

Cloning Scheme for Generating pCMV-NS35

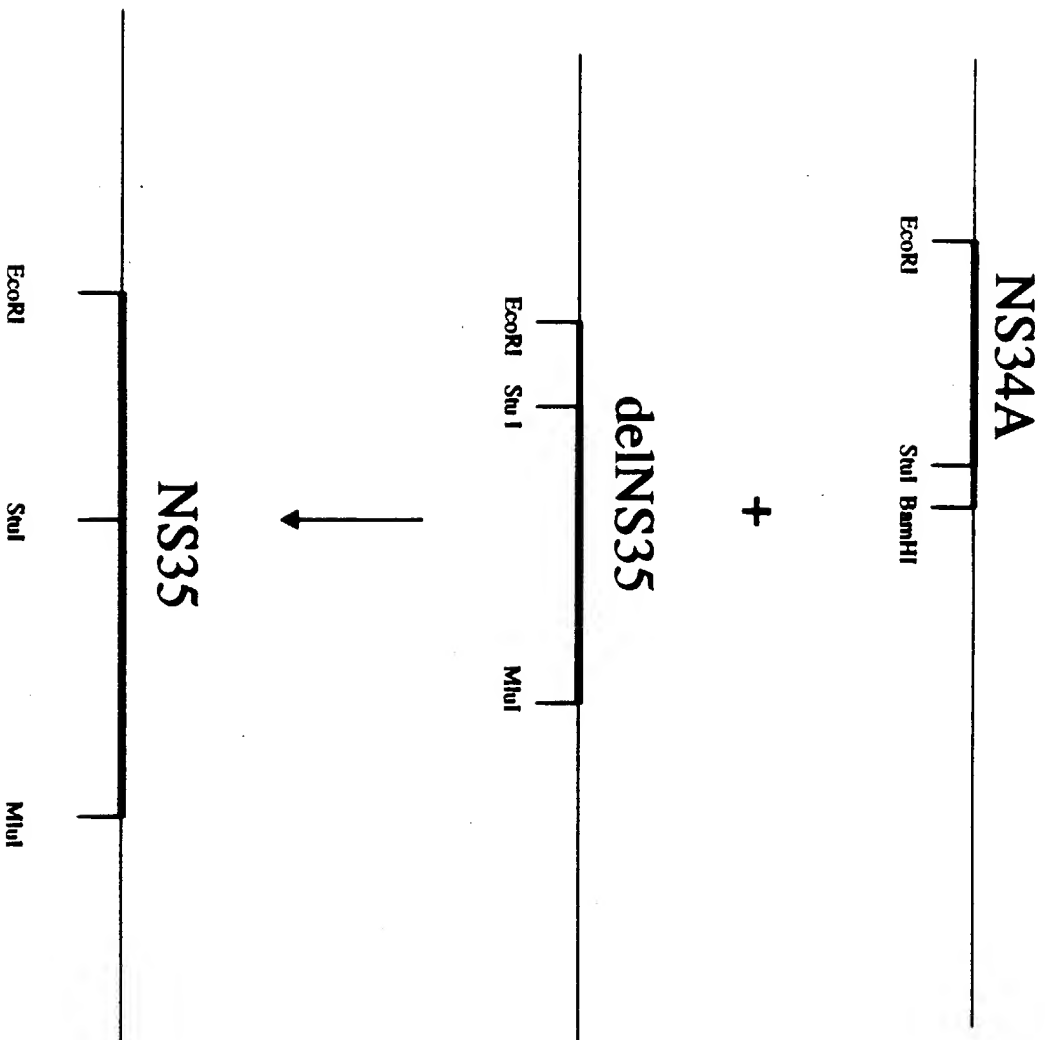


FIGURE 1

FIGURE 2

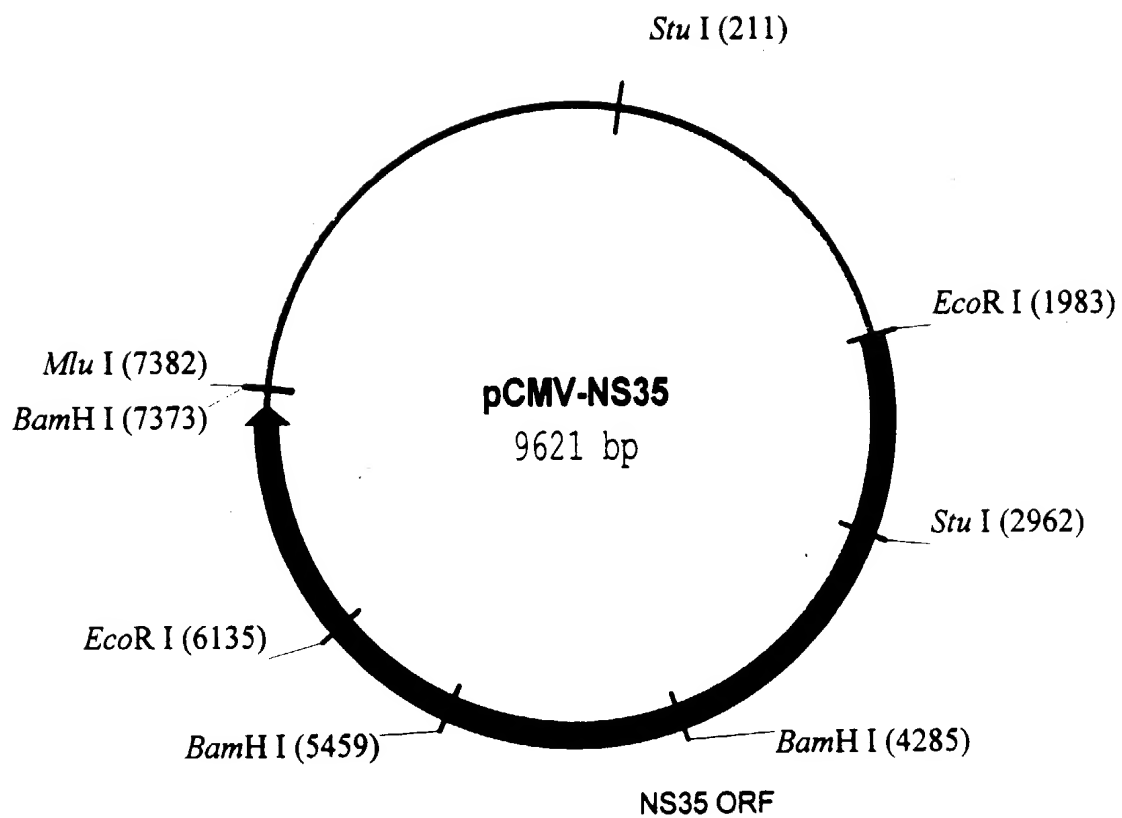


FIGURE 3 - Page 1

1	TCGCGCGTTT	CGGTGATGAC	GGTGAAAACC	TCTGACACAT	GCAGCTCCCG	GAGACGGTCA	CAGCTTGTCT	GTAAGCGGAT
	AGCGCGCAAA	GCCACTACTG	CCACTTTTGG	AGACTGTGTA	CGTCGAGGGC	CTCTGCCAGT	GTCGAACAGA	CATTGCGCTA
81	GCCGGGAGCA	GACAAGCCCG	TCAGGGCGCG	TCAGCGGGTG	TTGGCGGGTG	TCGGGGCTGG	CTTAACATATG	CGGCATCAGA
	CGGCCCTCGT	CTGTTCGGGC	AGTCCCGCGC	AGTCGCCCAC	AACCGCCCAC	AGCCCCGACC	GAATTGATAC	GCCGTAGTCT
					StuI	~~~~~		
161	GCAGATTGTA	CTGAGAGTGC	ACCATATGAA	GCTTTTTGCA	AAAGCCTAGG	CCTCCAAAAA	AGCCTCCTCA	CTACTTCTGG
	CGTCTAACAT	GACTCTCACG	TGGTATACTT	CGAAAAACGT	TTTCGGATCC	GGAGGTTTTT	TCGGAGGAGT	GATGAAGACC
241	AATAGCTCAG	AGGCCGAGGC	GGCCTCGGCC	TCTGCATAAA	TAAAAAAAT	TAGTCAGCCA	TGGGGCGGAG	AATGGGCGGA
	TTATCGAGTC	TCCGGGTCCG	CCGGAGCCGG	AGACGTATTT	ATTTTTTTTA	ATCAGTCGGT	ACCCCGCCTC	TTACCCGCCT
321	ACTGGGCGGG	GAGGGAATTA	TTGGCTATTG	GCCATTGCAT	ACGTTGTATC	TATATCATAA	TATGTACATT	TATATTGGCT
	TGACCCGCC	CTCCCTTAAT	AACCGATAAC	CGGTAACGTA	TGCAACATAG	ATATAGTATT	ATACATGTAA	ATATAACCGA
401	CATGTCCAAT	ATGACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA	TAGTAATCAA	TTACGGGGTC	ATTAGTTCAT
	GTACAGGTTA	TACTGGCGGT	ACAACTGTAA	CTAATAACTG	ATCAATAATT	ATCATTAGTT	AATGCCCCAG	TAATCAAGTA
481	AGCCCATATA	TGGAGTTCCG	CGTTACATAA	CTTACGGTAA	ATGGCCCGCC	TGGCTGACCG	CCCAACGACC	CCCGCCCAT
	TCGGGTATAT	ACCTCAAGGC	GCAATGTATT	GAATGCCATT	TACCGGGCGG	ACCGACTGGC	GGGTGTCTGG	GGGCGGGTAA
561	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AACGCCAATA	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGGT
	CTGCAGTTAT	TACTGCATAC	AAGGGTATCA	TTGCGGTTAT	CCCTGAAAGG	TAAGTGCAGT	TACCCACCTC	ATAAATGCCA
641	AAACTGCCCC	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTCCGCCC	CCTATTGACG	TCAATGACGG	TAAATGGCCC
	TTTGACGGGT	GAACCGTCAT	GTAGTTCACA	TAGTATACGG	TTCAGGCGGG	GGATAACTGC	AGTTACTGCC	ATTTACCGGG
721	GCCTGGCATT	ATGCCCAGTA	CATGACCTTA	CGGGACTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TCGCTATTAC
	CGGACCGTAA	TACGGGTCAT	GTACTGGAAT	GCCCTGAAAG	GATGAACCGT	CATGTAGATG	CATAATCAGT	AGCGATAATG
801	CATGGTGATG	CGGTTTTTGGC	AGTACACCAA	TGGGCGTGGA	TAGCGGTTTG	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA
	GTACCACTAC	GCCAAAACCG	TCATGTGGTT	ACCCGCACCT	ATCGCCAAAC	TGAGTGCCCC	TAAAGGTTCA	GAGGTGGGGT
881	TTGACGTCAA	TGGGAGTTTG	TTTTGGCACC	AAAATCAACG	GGACTTTCCA	AAATGTCGTA	ATAACCCCGC	CCCGTTGACG
	AACTGCAGTT	ACCCTCAAAC	AAAACCGTGG	TTTTAGTTGC	CCTGAAAGGT	TTTACAGCAT	TATTGGGGCG	GGGCAACTGC
961	CAAATGGGCG	GTAGGCGTGT	ACGGTGGGAG	GTCTATATAA	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG
	GTTTACCCGC	CATCCGCACA	TGCCACCCTC	CAGATATATT	CGTCTCGAGC	AAATCACTTG	GCAGTCTAGC	GGACCTCTGC
1041	CCATCCACGC	TGTTTTGACC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGCCCG	GGAACGGTGC	ATTGGAACGC
	GGTAGGTGCG	ACAAAACCTG	AGGTATCTTC	TGTGGCCCTG	GCTAGGTCCG	AGGCGCCGGC	CCTTGCCACG	TAACCTTGCG
1121	GGATTCCCCG	TGCCAAGAGT	GACGTAAGTA	CCGCCTATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA
	CCTAAGGGGC	ACGGTTCTCA	CTGCATTCT	GGCGGATATC	TGAGATATCC	GTGTGGGGAA	ACCGAGAATA	CGTACGATAT
1201	CTGTTTTTGG	CTTGGGGCCT	ATACACCCCC	GCTCCTTATG	CTATAGGTGA	TGGTATAGCT	TAGCCTATAG	GTGTGGGTAT
	GACAAAAACC	GAACCCCGGA	TATGTGGGGG	CGAGGAATAC	GATATCCACT	ACCATATCGA	ATCGGATATC	CACACCCAAT
1281	TTGACCATTA	TTGACCACTC	CCCTATTGGT	GACGATACTT	TCCATTACTA	ATCCATAACA	TGGCTCTTTG	CCACAACAT
	AACTGGTAAT	AACTGGTGAG	GGGATAACCA	CTGCTATGAA	AGGTAATGAT	TAGGTATTGT	ACCGAGAAAC	GGTGTGATA
1361	CTCTATTGGC	TATATGCCAA	TACTCTGTCC	TTCAGAGACT	GACACGGACT	CTGTATTTTT	ACAGGATGGG	GTCCATTTAT
	GAGATAACCG	ATATACGGTT	ATGAGACAGG	AAGTCTCTGA	CTGTGCCTGA	GACATAAAAA	TGTCCTACCC	CAGGTAAATA

002449.1200

002449.1200

FIGURE 3 - Page 2

1441 TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
ATAAATGTTT AAGTGTATAT GTTGTTCGGG CAGGGGGCAC GGGCGTCAAA AATAATTGT ATCGCACCTT AGAGGCTGTA

1521 CTCGGGTACG TGTTCCGGAC ATGGGCTCTT CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
GAGCCCATGC ACAAGGCCGTG TACCCGAGAA GAGGCCATCG CCGCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTTACGG GTGGTGGTGG

1681 AGTGTGCCGC ACAAGGCCGT GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT
TCACACGGCG TGTTCCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA CCTGCGTCTA

1761 GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT
CCTTCTGAAT TCCGTCGCCG TCTTCTCTA CGTCCGTCGA CTCACAAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841 TCGGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT
ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

+2 EcoRI M A A

1921 GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCTGCAT
CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGACGTA

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA ACACTGGGCT TTGGTGCTTA CATGTCCAAG
TACGTCGAGT CCCGATATTC CACGATCATG AGTTGGGGAG ACAACGACGT TGTGACCCGA AACCACGAAT GTACAGGTTT

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G
2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCAGTACT CCACCTACGG
CGAGTACCCT AGCTAGGATT GTAGTCTG GCGGCTCTT GTTAATGGTG ACCGTCGGGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A
2161 CAAGTTCCTT GCCGACGGCG GGTGCTCGGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA
GTTCAAGGAA CGGCTGCCGC CCACGAGCCC CCGCGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T
2241 CATCCATCTT GGGCATTGGC ACTGTCCTTG ACCAAGCAGA GACTGCGGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC
GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGTTTCGTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G
2321 CCTCCGGGCT CCGTCACTGT GCGCCATCCC AACATCGAGG AGGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTACGG
GGAGGCCCGA GGCAGTGACA CGGGGTAGGG TTGTAGCTCC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L
2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCATTCAAA GAAGAAGTGC GACGAAGTGC
GTTCCGATAG GGGGAGCTTC ATTAGTCCC CCCCTCTGTA GAGTAGAAGA CAGTAAGTTT CTGTTTACG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G
2481 CCGCAAAGCT GGTGCGATTG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC
GGCGTTTCGA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGCGCCA GAACTGCACA GGCAGTAGGG CTGGTGGCGG

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C
2561 GATGTTGTCG TCGTGGCAAC CGATGCCCTC ATGACCGGCT ATACCGGCGA CTTGACTCG GTGATAGACT GCAATACGTG
CTACAACAGC AGCACCCTTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC

097479 1524260

FIGURE 3 - Page 3

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S
 2641 . TGTCACCCAG ACAGTCGATT TCAGCCTTGA CCCTACCTTC ACCATTGAGA CAATCACGCT CCCCCAAGAT GCTGTCTCCC
 ACAGTGGGTC TGTCAGCTAA AGTCGGAAGT GGGATGGAAG TGGTAACTCT GTTAGTGCGA GGGGGTTCTA CGACAGAGGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G
 2721 GCACTCAACG TCGGGGAGG ACTGGCAGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGGC
 CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AAACACCGTG GCCCCCTCGC GGGGAGGCGC

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V
 2801 ATGTTCTGACT CGTCCGTCTT CTGTGAGTGC TATGACGCAG GCTGTGCTTG GTATGAGCTC ACGCCCGCCG AGACTACAGT
 TACAAGCTGA GCAGGCAGGA GAACTCACG AACTGCGTC CGACACGAAC CATACTCGAG TGCGGGCGGC TCTGATGTCA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T
 2881 TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG
 ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGGTCCTGG TAGAACTTAA AACCCCTCCG CAGAAATGTC

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q
 StuI
 2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA
 CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCACCC CTCTTGGAAG GAATGGACCA TCGCATGGTT

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T
 3041 GCCACCGTGT GCGCTAGGGC TCAAGCCCTT CCCCCATCGT GGGACCAGAT GTGGAAGTGT TTGATTCGCC TCAAGCCAC
 CGGTGGCACA CGCGATCCCG AGTTCGGGGA GGGGGTAGCA CCCTGGTCTA CACCTTCACA AACTAAGCGG AGTTCGGGTG

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K
 3121 CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGGCGCT GTTCAGAATG AAATCACCTT GACGCACCCA GTCACCAAAT
 GGAGTACCC GGTGTGGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTTAGTGGGA CTGCGTGGGT CAGTGGTTTA

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L
 3201 ACATCATGAC ATGCATGTCG GCCGACCTGG AGGTCGTCAC GAGCACCTGG GTGCTCGTTG GCGGCGTCTT GGCTGCTTTG
 TGTAGTACTG TACGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGAGGA CCGACGAAAC

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D
 3281 GCCGCGTATT GCCTGTCAAC AGGCTGCGTG GTCATAGTGG GCAGGGTCTG CTTGTCCGGG AAGCCGGCAA TCATACCTGA
 CGGCGCATAA CGGACAGTTG TCCGACGCAC CAGTATCACC CGTCCCAGCA GAACAGGCCC TTCGGCCGTT AGTATGGACT

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M
 3361 CAGGGAAGTC CTCTACCGAG AGTTCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC
 GTCCTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCTACTACG

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V
 3441 TCGCCGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGCGTCC CGTCAGGCAG AGGTTATCGC CCCTGCTGTC
 AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G
 3521 CAGACCAACT GGCAAAACT CGAGACCTTC TGGGCGAAGC ATATGTGGAA CTTATCAGT GGGATACAAT ACTTGGCGGG
 GTCTGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTCG TATACACCTT GAAGTAGTCA CCCTATGTTA TGAACCGCCC

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T
 3601 CTTGTCAACG CTGCCTGGTA ACCCGGCCAT TGCTTCATTG ATGGCTTTTA CAGCTGCTGT CACCAGCCCA CTAACCACTA
 GAACAGTTGC GACGGACCAT TGGGGCGGTA ACGAAGTAAC TACCGAAAAT GTCGACGACA GTGGTGGGT GATTGGTGT

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V
 3681 GCCAAACCTT CCTCTCAAC ATATTGGGGG GGTGGGTGGC TGCCAGCTC GCCGCCCCCG GTGCCGCTAC TGCCTTTGTG
 CGGTTTGGGA GGAGAAGTTG TATAACCCCC CCACCCACCG ACGGGTCGAG CGGCGGGGGC CACGGCGATG ACGGAAACAC

FIGURE 3 - Page 4

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A
 3761 G G C G T G G C T T A G C T G G C G C G C C A T C G G C A G T G T T G G A C T G G G A A G G T C C T C A T A G A C A T C C T T G C A G G G T A T G G C G A
 C C G C G A C C G A A T C G A C C G C G C G G T A G C C G T C A C A A C C T G A C C C C T T C C A G G A T A T C T G T A G G A A C G T C C C A T A C C G C G

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L
 3841 G G C G T G G C G G G A G C T C T T G T G C A T T C A A G A T C A T G A G C G G T G A G G T C C C T C C A C G G A G G A C C T G G T C A A T C T A C T G C
 C C C G C A C C G C C C T C G A G A A C A C C G T A A G T T C T A G T A C T C G C C A C T C C A G G G A G G T G C C T C T G G A C C A G T T A G A T G A C G

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E
 3921 C C G C C A T C C T C T C G C C C G G A G C C C T C G T A G T C G G C G T G G T C T G T G C A G C A A T A C T G C G C C G G C A C G T T G C C C G G G C G A G
 G G C G T A G G A G A G C G G G C C T C G G G A G C A T C A G C C G C A C C A G A C A C G T C G T T A T G A C G C G G C C G T G C A A C C G G C C C G C T C

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E
 4001 G G G G C A G T G C A G T G G A T G A A C C G G C T G A T A G C T T C G C C T C C C G G G G G A A C C A T G T T T C C C C A C G C A C T A C G T G C C G G A
 C C C C G T C A C G T C A C C T A C T T G G C C G A C T A T C G G A A G C G G A G G G C C C C T T G G T A C A A A G G G G T G C G T G A T G C A C G G C C T

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W
 4081 G A G C G A T G C A G C T G C C C G C G T A C T G C C A T A C T C A G C A G C C T C A C T G T A A C C A G C T C C T G A G G C G A C T G C A C C A G T G G A
 C T C G C T A C G T C G A C G G G C G C A G T G A C G G T A T G A G T C G T C G G A G T G A C A T T G G G T C G A G G A C T C C C G T G A C G T G G T C A C C T

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D
 4161 T A A G C T C G G A G T G T A C C A C T C C A T G C T C C G G T T C C T G G C T A A G G G A C A T C T G G G A C T G G A T A T G C G A G G T G T T G A G C G A C
 A T T C G A G C C T C A C A T G G T G A G G T A C G A G G C C A A G G A C C G A T T C C C T G T A G A C C C T G A C C T A T A C G C T C C A C A A C T C G C T G

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G
 BamHI
 ~~~~~  
 4241 T T T A A G A C C T G G C T A A A A G C T A A G C T C A T G C C A C A G C T G C T G G G A T C C C C T T T G T G T C C T G C C A G C G C G G G T A T A A G G G  
 A A A T T C T G G A C C G A T T T T C G A T T C G A G T A C G G T G T C G A C G G A C C C T A G G G G A A C A C A G G A C G G T C G C G C C A T A T T C C C

---

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T  
 4321 G G T C T G G C G A G G G G A C G G C A T C A T G C A C A C T C G C T G C C A C T G T G G A G C T G A G A T C A C T G G A C A T G T C A A A A A C G G G G A C G A  
 C C A G A C C G C T C C C C T G C C G T A G T A C G T G T G A G C G A C G G T G A C A C C C T C G A C T C T A G T G A C C T G T A C A G T T T T G C C C T G C T

---

+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C  
 4401 T G A G G A T C G T C G G T C C T A G G A C C T G C A G G A A C A T G T G G A G T G G G A C C T T C C C A T T A A T G C C T A C A C C A C G G G C C C C T G T  
 A C T C C T A G C A G C C A G G A T C C T G G A C G T C C T G T A C A C C T C A C C T G G A A G G G G T A A T T A C G G A T G T G G T G C C C G G G G A C A

---

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G  
 4481 A C C C C C C T T C T G C G C C G A A C T A C A C G T T C G C G T A T G G A G G G T G T C T G C A G A G G A A T A C G T G G A G A T A A G G C A G G T G G G  
 T G G G G G A A G G A C G C G G C T T G A T G T G C A A G C G C G A T A C C T C C C A C A G A C G T C T C C T T A T G C A C C T C T A T T C C G T C C A C C C

---

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T  
 4561 G G A C T T C C A C T A C G T G A C G G G T A T G A C T A C T G A C A A T C T T A A T G C C C C G T G C C A G G T C C C A T C G C C C G A A T T T T T C A C A G  
 C C T G A A G G T G A T G C A C T G C C C A T A C T G A T G A C T G T T A G A A T T A C G G G C A C G G T C C A G G G T A G C G G G C T T A A A A A G T G T C

---

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G  
 4641 A A T T G G A C C G G G T G C G C C T A C A T A G G T T T G C G C C C C C T G C A A G C C C T T G C T G C G G G A G G A G G T A T C A T T C A G A G T A G G A  
 T T A A C C T G C C C A C G C G G A T G T A T C C A A A C G C G G G G G G A C G T T C G G G A A C G A C G C C C T C C T C C A T A G T A A G T C T A T C C T

---

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D  
 4721 C T C C A C G A A T A C C C G G T A G G G T C G C A A T T A C C T T G C G A G C C C G A A C C G A C G T G C C G T G T T G A C G T C C A T G C T A C T G A  
 G A G G T G C T T A T G G G C C A T C C A G C G T T A A T G G A A C G C T C G G G C T T G G C C T G C A C C G G C A C A A C T G C A G G T A C G A G T G A C T

---

+2 P S H I T A E A A G R R L A R G S P P S V A S S S A  
 4801 T C C C T C C C A T A T A A C A G C A G A G G C G G C C G G G C G A A G G T T G C G A G G G G A T C A C C C C C C T C T G T G G C C A G C T C C T C G G C T A  
 A G G G A G G G T A T A T T G T C G T C T C C G C C G G C C G C T T C C A A C C G T C C C C T A G T G G G G G A G A C A C C G G T C G A G G A G C C G A T

## FIGURE 3 - Page 5

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N  
 4881 GCCAGCTATC CGCTCCATCT CTAAGGCAA CTTGCACCGC TAACCATGAC TCCCCTGATG CTGAGCTCAT AGAGGCCAAC  
 CCGTTCGATAG GCGAGGTAGA GAGTTCGGT GAACGTGGCG ATTGGTACTG AGGGGACTAC GACTCGAGTA TCTCCGGTTG

---

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D  
 4961 CTCCTATGGA GGCAGGAGAT GGGCGGCAAC ATCACCAGGG TTAGTCAAGA AAACAAAGTG GTGATTCTGG ACTCCTTCGA  
 GAGGATACCT CCGTCTCTA CCCGCCGTTG TAGTGGTCCC AACTCAGTCT TTTGTTTCAC CACTAAGACC TGAGGAAGCT

---

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q  
 5041 TCCGCTTGTT GCGGAGGAGG ACGAGCGGGA GATCTCCGTA CCCGCAGAAA TCCTGCGGAA GTCTCGGAGA TTCGCCCAGG  
 AGGCGAACAC CGCTCTCTCC TGCTCGCCCT CTAGAGGCAT GGGCGTCTTT AGGACGCCTT CAGAGCCTCT AAGCGGGTCC

---

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V  
 5121 CCCTGCCCCG TTGGGCGCGG CCGGACTATA ACCCCCCGCT AGTGGAGACG TGGAAAAAGC CCGACTACGA ACCACCTGTG  
 GGGACGGGCA AACCCGCGCC GGCCTGATAT TGGGGGGCGA TCACCTCTGC ACCTTTTTTCG GGCTGATGCT TGGTGGACAC

---

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E  
 5201 GTCCATGGCT GCCCGCTTCC ACCTCCAAAG TCCCCTCCTG TGCCTCCGCC TCGGAAGAAG CCGACGGTGG TCCTCACTGA  
 CAGGTACCGA CGGGCGAAGG TGGAGGTTTC AGGGGAGGAC ACGGAGGCGG AGCCTTCTTC GCCTGCCACC AGGAGTGACT

---

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D  
 5281 ATCAACCCTA TCTACTGCCT TGGCCGAGCT CGCCACCAGA AGCTTTGGCA GCTCCTCAAC TTCCGGCATT ACGGGCGACA  
 TAGTTGGGAT AGATGACGGA ACCGGCTCGA GCGGTGGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTAA TGCCCGCTGT

---

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P  
 5361 ATACGACAAC ATCCTCTGAG CCCGCCCTT CTGGTGCCC CCCCAGTCC GACGCTGAGT CCTATTCTC CATGCCCCC  
 TATGCTGTTG TAGGAGACTC GGGCGGGGAA GACCGACGGG GGGGCTGAGG CTGCGACTCA GGATAAGGAG GTACGGGGG

---

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V  
 ~~~~~  
 BamHI

5441 CTGGAGGGGG AGCCTGGGGA TCCGATCTT AGCGACGGGT CATGGTCAAC GGTCAAGTAGT GAGGCCAACG CGGAGGATGT
 GACCTCCCCC TCGGACCCCT AGGCCTAGAA TCGTGCCCA GTACCAGTTG CCAGTCATCA CTCCGGTTGC GCCTCTTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I
 5521 CGTGTGCTGC TCAATGTCTT ACTCTGGAC AGGCGCACTC GTCACCCCGT GCGCCGCGGA AGAACAGAAA CTGCCCATCA
 GCACACGACG AGTTACAGAA TGAGAACCTG TCCGCGTGAG CAGTGGGGCA CGCGGCGCCT TCTTGTCTTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K
 5601 ATGCACTAAG CAACTCGTTG CTACGTCAAC ACAATTGGT GTATTCCACC ACCTCAGCA GTGCTTGCCA AAGGCAGAA
 TACGTGATTC GTTGAGCAAC GATGCAGTGG TGTTAAACCA CATAAGGTG TGGAGTGCCT CACGAACGGT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A A S K
 5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA
 TTTCAGTGTA AACTGTCTGA CGTTCAAGAC CTGTCGGTAA TGGTCTGCA TGAGTTCCTC CAATTCGTC GCCGAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y
 5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGACGC CTGACGCCCC CACACTCAGC CAAATCCAAG TTTGGTTATG
 TCACTTCCGA TTGAACGATA GGCATCTCCT TCGAACGTCG GACTGCGGGG GTGTGAGTCG GTTTAGGTTT AAACCAATAC

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N
 5841 GGGCAAAAGA CGTCCGTTGC CATGCCAGAA AGGCCGTAAC CCACATCAAC TCCGTGTGGA AAGACCTTCT GGAAGACAAT
 CCCGTTTTCT GCAGGCAACG GTACGGTCTT TCCGGCATTG GGTGTAGTTG AGGCACACCT TTCTGGAAGA CCTTCTGTTA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A
 5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCCTGAG AAGGGGGGTC GTAAGCCAGC
 CATTGTGGTT ATCTGTGATG GTAGTACCGA TTCTGTCTCC AAAAGACGCA AGTCGGACTC TTCCCCCAG CATTGCGTCC

+2	R L I V F P D L G V R V C E K M A L Y D V V T K L P	
6001	TCGTCTCATC GTGTTCCCCG ATCTGGGCGT GCGCGTGTGC GAAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCCT AGCAGAGTAG CACAAGGGGC TAGACCCGCA CGCGCACACG CTTTTCTACC GAAACATGCT GCACCAATGT TTCGAGGGGA	
+2	L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S	EcoRI
6081	TGGCCGTGAT GGAAGCTCC TACGGATTCC AATACTCACC AGGACAGCGG GTTGAATTCC TCGTGCAAGC GTGGAAGTCC ACCGGCACTA CCCTTCGAGG ATGCCTAAGG TTATGAGTGG TCCTGTGCGC CAACTTAAGG AGCACGTTCC CACCTTCAGG	
+2	K K T P M G F S Y D T R C F D S T V T E S D I R T E E	
6161	AAGAAAACCC CAATGGGGTT CTCGTATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGCGACATCC GTACGGAGGA TTCTTTTGGG GTTACCCCAA GAGCATACTA TGGGCGACGA AACTGAGGTG TCAGTGACTC TCCTGTAGG CATGCTCTCT	
+2	A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G	
6241	GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTTCCGGCG CACCGGTAGT TCAGGGAGTG GCTCTCCGAA ATACAAACCC	
+2	G P L T N S R G E N C G Y R R C R A S G V L T T S C G	
6321	GCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGCTATCG CAGGTGCCGC GCGAGCGGCG TACTGACAAC TAGCTGTGGT CGGGAGAATG GTTAAGTTCC CCCCCTTTGA CGCCGATAGC GTCCACGGCG CGCTCGCCGC ATGACTGTTG ATCGACACCA	
+2	N T L T C Y I K A R A A C R A A G L Q D C T M L V C G	
6401	AACACCCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCGAGG GCTCCAGGAC TGCACCATGC TCGTGTGTGG TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCTCG ACGTGGTACG AGCACACACC	
+2	D D L V V I C E S A G V Q E D A A S L R A F T E A M	
6481	CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGGTCCAG GAGGACGCGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA GCTGCTGAAT CAGCAATAGA CACTTTCGCG CCCCAGGTG CTCCTGCGCC GCTCGGACTC TCGGAAGTGC CTCCGATACT	
+2	T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V	
6561	CCAGGTACTC CGCCCCCCT GGGGACCCCC CACAACCAGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTG GGTCCATGAG GCGGGGGGGA CCCCTGGGGG GTGTTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC	
+2	S V A H D G A G K R V Y Y L T R D P T T P L A R A A W	
6641	TCAGTCGCCC ACGACGGCGC TGGAAAGAGG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGCGA GAGCTGCGTG AGTCAGCGGG TGCTGCCGCG ACCTTTCTCC CAGATGATGG AGTGGGCACT GGGATGTTGG GGGGAGCGCT CTCGACGCAC	
+2	E T A R H T P V N S W L G N I I M F A P T L W A R M	
6721	GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTTGCCCC CACACTGTGG GCGAGGATGA CCTCTGTGCT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAAACGGGG GTGTGACACC CGCTCCTACT	
+2	I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A	
6801	TACTGATGAC CCATTTCTTT AGCGTCCTTA TAGCCAGGGA CCAGCTTGAA CAGGCCCTCG ATTGCGAGAT CTACGGGGCC ATGACTACTG GGTAAGAAA TCGCAGGAAT ATCGGTCCCT GGTCGAACCT GTCCGGGAGC TAACGCTCTA GATGCCCCGG	
+2	C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y	
6881	TGCTACTCCA TAGAACCACT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA ACGATGAGGT ATCTTGGTGA CCTAGATGGA GGTTAGTAAG TTTCTGAGGT ACCGGAGTCG CGTAAAAGTG AGGTGTCAAT	
+2	S P G E I N R V A A C L R K L G V P P L R A W R H R	
6961	CTCTCCAGGT GAAATCAATA GGGTGGCCGC ATGCCTCAGA AAACCTTGGGG TACCGCCCTT GCGAGCTTGG AGACACCGGG GAGAGGTCCA CTTTAGTTAT CCCACCGGCG TACGGAGTCT TTTGAACCC ATGGCGGGAA CGCTCGAACC TCTGTGGCCC	
+2	A R S V R A R L L A R G G R A A I C G K Y L F N W A V	
7041	CCCGGAGCGT CCGCGCTAGG CTTCTGGCCA GAGGAGGCAG GGCTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA GGGCCTCGCA GGCGCGATCC GAAGACCGGT CTCCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT	

FIGURE 3 - Page 7

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G
 7121 AGAACAAAGC TCAAATCAC TCCAATAGCG GCCGCTGGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG
 TCTTGTTCG AGTTTGAGTG AGGTTATCGC CGGCGACCGG TCGACCTGAA CAGGCCGACC AAGTGCCGAC CGATGTCGCG

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V
 7201 GGGAGACATT TATCACAGCG TGTCTCATGC CCGGCCCGCG TGGATCTGGT TTTGCCTACT CCTGCTTGCT GCAGGGGTAG
 CCCTCTGTAA ATAGTGTGCG ACAGAGTACG GGCCGGGGCG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCCATC

+2 G I Y L L P N R
 7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAAAAA AAAAATCTAG AAAGGCGCGC
 CGTAGATGGA GGAGGGGTG GCTACTTCCA ACCCCATTG TGAGGCCGGA TTTTTTTTTT TTTTATAGT TTTCCGCGCG

BamHI MluI
 7361 CAAGATATCA AGGATCCACT ACGCGTTAGA GCTCGCTGAT CAGCCTCGAC TGTGCCTTCT AGTTGCCAGC CATCTGTTGT
 GTTCTATAGT TCCTAGGTGA TCGCAATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA TCAACGGTCG GTAGACAACA

7441 TTGCCCTCC CCGTGCCTT CCTTGACCCT GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAAATGAG GAAATTGCAT
 AACGGGGAGG GGGCAGCGAA GGAAGTGGGA CCTTCCACGG TGAGGGTGAC AGGAAAGGAT TATTTTACTC CTTTAACGTA

7521 CGCATTGTCT GAGTAGGTGT CATTCTATTC TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT
 GCGTAACAGA CTATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTCGTTCC CCCTCCTAAC CTTTCTGTTA

7601 AGCAGGCATG CTGGGGAGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT
 TCGTCCGTAC GACCCCTCGA GAAGGCGAAG GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCCGACGCC GCTCGCCATA

7681 CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG
 GTCGAGTGAG TTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCTATTG CGTCTTTTCT TGTACACTCG TTTTCCGGTC

7761 CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA
 GTTTTCCGGT CCTTGGCATT TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCGGGG GGACTGCTCG TAGTGTTTTT

7841 TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATAACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC
 AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCTGATG ATTTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCAGC

7921 GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC
 CGAGAGGACA AGGCTGGGAC GGCGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACC CGCA AAGAGTTACG

8001 TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA
 AGTGGGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTTCGACCC GACACACGTG CTTGGGGGGC AAGTCGGGCT

8081 CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG
 GGCGACGCGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CCGTGACCGT CGTCGGTGAC

8161 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACACGG CTACACTAGA
 CATTGTCTTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT

8241 AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAACA
 TCCTGTCATA AACCATAGAC GCGAGACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA TCGAGAATA GGCCGTTTGT

8321 AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT
 TTGGTGGCGA CCATCGCCAC CAAAAAACA AACGTTGCTG GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTTCTAGGAA

8401 TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG
 ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA TAGTTTTTCC

FIGURE 3 - Page 8

8481 ATCTTCACCT AGATCCTTTT AAATTA AAAA TGAAGTTT AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG
TAGAAGTGGA TCTAGGAAAA TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA CCAGACTGTC

8561 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTT GTTCATCCAT AGTTGCCTGA CTCCCCGTGG
AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCGAG

8641 TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT
ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACGGGG GTCACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA

8721 CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA
GGTCTAAATA GTCGTTATTT GGTGGGTGCG CCTTCCCGGC TCGCGTCTTC ACCAGGACGT TGAAATAGGC GGAGGTAGGT

8801 GTCTATTAAT TGTGCGGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG
CAGATAATTA ACAACGGCCC TTCGATCTCA TTCATCAAGC GGTCAATTAT CAAACGCGTT GCAACAACGG TAACGATGTC

8881 GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC
CGTAGCACCA CAGTGCGAGC AGCAAACCAT ACCGAAGTAA GTCGAGGCCA AGGGTTGCTA GTTCCGCTCA ATGTACTAGG

8961 CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTCGGTCTCT CGATCGTTGT CAGAAGTAAG TTGGCCGCGAG TGTATCACT
GGGTACAACA CGTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATTC AACC GGCGTC ACAATAGTGA

9041 CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCTAT CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA
GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAAG ACACTGACCA CTCATGAGTT

9121 CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT
GGTTCACTAA GACTCTTATC ACATACGCCG CTGGCTCAAC GAGAACGGGC CGCAGTTATG CCCTATTATG GCGCGGTGTA

9201 AGCAGAACTT TAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTGAGATC
TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTTGCAAGAA GCCCGGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

9281 CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA
GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTTT

9361 CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT
GTCCTTCGGT TTTACGGCGT TTTTCCCTT ATTCCCGCTG TGCCTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

9441 TATTGAAGCA TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC AAATAGGGGT
ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAATC TTTTATTG TTTATCCCCA

9521 TCCGCGCACA TTTCCCCGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGC
AGGCGCGTGT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTTGTAAT AATAGTACTG TAATTGGATA TTTTATCCG

9601 GTATCACGAG GCCCTTTCGT C
CATAGTGCTC CGGGAAAGCA G

8481
8561
8641
8721
8801
8881
8961
9041
9121
9201
9281
9361
9441
9521
9601

FIGURE 4

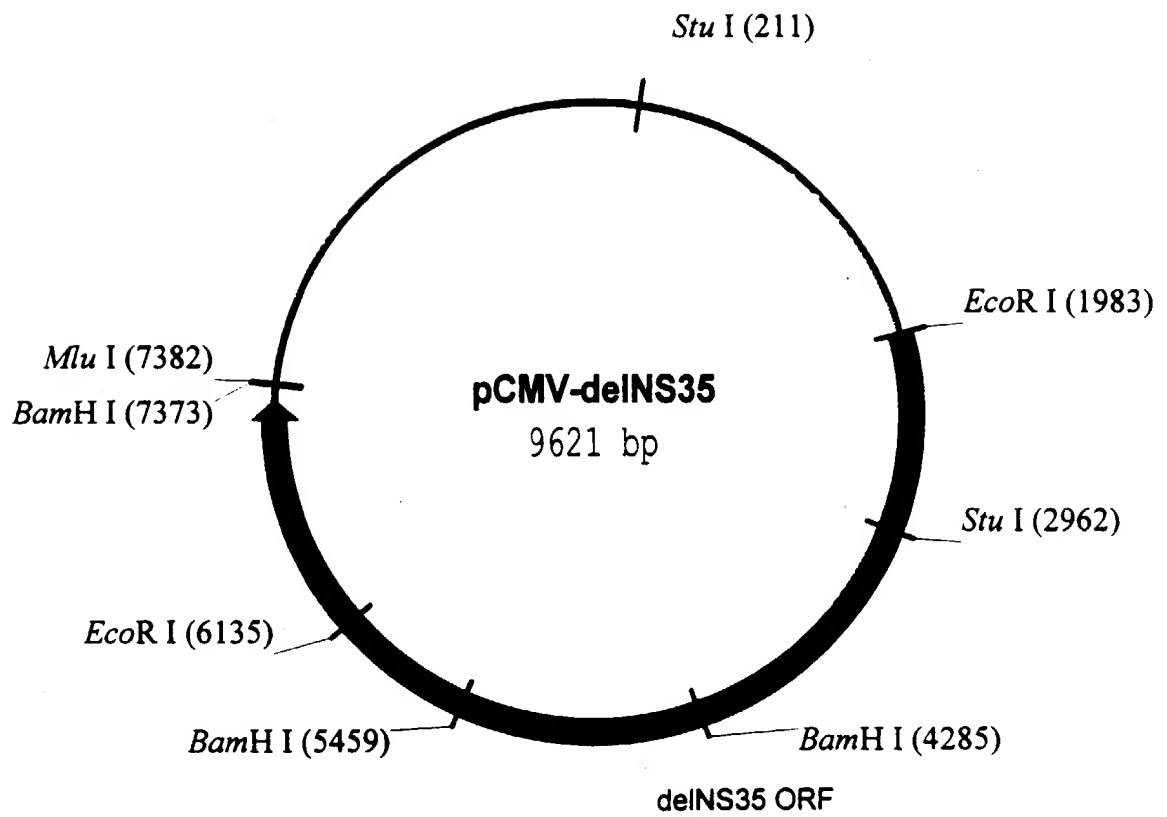


FIGURE 5 - Page 1

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT
AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTCCGCTA

81 GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA
CGGCCCTCGT CTGTTCCGGC AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT

StuI
~~~~~

161 GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTCGA AAAGCCTAGG CCTCCAAAA AGCCTCCTCA CTAATTCTGG  
CGTCTAACAT GACTCTCAG TGTATACTT CGAAAAACGT TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC

---

241 AATAGCTCAG AGGCCGAGGC GGCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGGCGGA  
TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT ACCCGCCTC TTACCCGCT

---

321 ACTGGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
TGACCCGCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

---

401 CATGTCCAAT ATGACCGCCA TGTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGT AATGCCCCAG TAATCAAGTA

---

481 AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCATT  
TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA

---

561 GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT  
CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

---

641 AACTGCCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCC CCTATTGACG TCAATGACGG TAAATGGCCC  
TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC AGTTACTGCC ATTTACGGG

---

721 GCCTGGCATT ATGCCAGTA CATGACCTTA CGGGACTTTC CTAATTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
CGGACCGTAA TACGGGTCAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

---

801 CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA  
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT

---

881 TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC CCGTTGACG  
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

---

961 CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG  
GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

---

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG GGAACGGTGC ATTGGAACGC  
GGTAGGTGCG ACAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC CCTTGCCACG TAACCTTGCG

---

1121 GGATTCCCCG TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
CCTAAGGGGC ACGGTTCTCA CTGCATTCTT GCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

---

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCG GTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA  
GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCAAT

---

1281 TTGACCATTA TTGACCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACAT  
AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

---

1361 CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTAT  
GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA

003377 5442650



## FIGURE 5 - Page 2

1441 TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA TAGCGTGGA TCTCCGACAT  
ATAAATGTTT AAGTGTATAT GTTGTTCGGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT ATCGCACCT AGAGGCTGTA

---

1521 CTCGGGTACG TGTTCGGAC ATGGGCTCTT CTCCGGTAGC GCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
GAGCCCATGC ACAAGGCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

---

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTAGGCACA GCACAATGCC CACCACCACC  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACGG GTGGTGGTGG

---

1681 AGTGTGCCGC ACAAGGCCGT GCGGCTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT  
TCACACGGCG TGTTCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGA CCTGCGTCTA

---

1761 GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT  
CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

---

1841 TGCGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT  
ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

---

+2 M A A  
EcoRI  
~~~~~

1921 GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCTGCAT
CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGCCA GCAGCTGGAT TCTTAAGTGG TACCGACGTA

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K

2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA AACTGGGCT TTGGTGCTTA CATGTCCAAG
TACGTCGAGT CCCGATATTC CACGATCATG AGTTGGGGAG ACAACGACGT TGTGACCCGA AACCACGAAT GTACAGGTTC

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G

2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCACGTAAT CCACCTACGG
CGAGTACCCT AGCTAGGATT GTAGTCTTGG CCCACTCTT GTTAATGGTG ACCGTCGGGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A

2161 CAAGTTCCTT GCCGACGGCG GGTGCTCGGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA
GTTCAAGGAA CGGCTGCCGC CCACGAGCCC CCCGCGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T

2241 CATCATCTT GGGCATTGGC ACTGTCCTTG ACCAAGCAGA GACTGCGGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC
GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTTCGTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G

2321 CCTCCGGGCT CCGTCACTGT GCGCCATCCC AACATCGAGG AGGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTACGG
GGAGGCCCCG GGCAGTGACA CGGGGTAGGG TTGTAGTCTC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L

2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCATTCAAA GAAGAAGTGC GACGAAGTGC
GTTCCGATAG GGGGAGCTTC ATTAGTTCCC CCCCTCTGTA GAGTAGAAGA CAGTAAGTTT CTCTTTCACG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G

2481 CCGCAAAGCT GGTGCGATTG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC
GGCGTTTCGA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGCGCCA GAACTGCACA GGCAGTAGGG CTGGTGGCGG

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C

2561 GATGTTGTCG TCGTGGAAC CGATGCCCTC ATGACCGGCT ATACCGGCGA CTTGACTCG GTGATAGACT GCAATACGTG
CTACAACAGC AGCACCCTTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC

FIGURE 5 - Page 3

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S
 2641 TGTCACCCAG ACAGTCGATT TCAGCCTTGA CCCTACCTTC ACCATTGAGA CAATCACGCT CCCCCAAGAT GCTGTCTCCC
 ACAGTGGGTC TGTCAGCTAA AGTCGGAAGT GGGATGGAAG TGGTAACTCT GTTAGTGCGA GGGGGTTCTA CGACAGAGGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G
 2721 GCACTCAACG TCGGGGCAGG ACTGGCAGGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGCG
 CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AAACACCGTG GCCCCCTCGC GGGGAGGCGG

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V
 2801 ATGTTTCTGACT CGTCCGTCCT CTGTGAGTGC TATGACGCAG GCTGTGCTTG GTATGAGCTC ACGCCCGCCG AGACTACAGT
 TACAAGCTGA GCAGGCAGGA GACACTCAG ATACTGCGTC CGACACGAAC CATACTCGAG TCGGGGCGGC TCTGATGTCA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T
 2881 TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG
 ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGGTCTCTG TAGAACTTAA AACCTCCCCG CAGAAATGTC

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q
 StuI
 2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA
 CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCACCC CTCTTGAAG GAATGGACCA TCGCATGTTT

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T
 3041 GCCACCGTGT GCGCTAGGGC TCAAGCCCCT CCCCCATCGT GGGACCAGAT GTGGAAGTGT TTGATTGCGC TCAAGCCCAC
 CGGTGGCACA CGCGATCCCG AGTTCGGGGA GGGGGTAGCA CCCTGGTCTA CACCTTACA AACTAAGCGG AGTTCGGGTG

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K
 3121 CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGGCGCT GTTCAGAATG AAATCACCTC GACGCACCCA GTCACCAAA
 GGAGGTACCC GGTTGTGGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTTAGTGGGA CTGCGTGGGT CAGTGGTTTA

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L
 3201 ACATCATGAC ATGCATGTCG GCCGACCTGG AGGTCGTCAC GAGCACCTGG GTGCTCGTTG GCGGCGTCTT GGCTGCTTTG
 TGTAAGTACTG TACGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGACGGA CCGACGAAAC

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D
 3281 GCCGCGTATT GCCTGTCAAC AGGCTGCGTG GTCATAGTGG GCAGGTCGT CTTGTCCGGG AAGCCGGCAA TCATACCTGA
 CGGCGCATAA CGGACAGTTG TCCGACGCAC CAGTATCACC CGTCCCAGCA GAACAGGCCC TTCGGCCGTT AGTATGGACT

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M
 3361 CAGGGAAGTC CTCTACCGAG AGTTCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC
 GTCCCTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCCTACTACG

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V
 3441 TCGCCGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGCGTCC CGTCAGGCAG AGGTTATCGC CCCTGCTGTC
 AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G
 3521 CAGACCAACT GGCAAAAACCT CGAGACCTTC TGGGCGAAGC ATATGTGGAA CTTTCATCAGT GGGATACAAT ACTTGGCGGG
 GTCTGGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTCG TATACACCTT GAAGTAGTCA CCCTATGTTA TGAACCGCCC

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T
 3601 CTTGTCAACG CTGCTGGTA ACCCGCCAT TGCTTCATTG ATGGCTTTTA CAGCTGCTGT CACCAGCCCA CTAACCTA
 GAACAGTTGC GACGGACCAT TGGGGCGGTA ACGAAGTAAC TACCGAAAAT GTCGACGACA GTGGTGGGT .GATTGGTGT

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V
 3681 GCCAAACCCT CCTCTTCAAC ATATTGGGGG GGTGGGTGGC TGCCAGCTC GCCGCCCCCG GTGCGGCTAC TGCCTTTGTG
 CGGTTTGGGA GGAGAAGTTG TATAACCCCC CCACCCACCG ACGGTTCGAG CGGCGGGGGC CACGGCGATG ACGGAAACAC

FIGURE 5 - Page 4

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A
 3761. GGCGCTGGCT TAGCTGGCGC CGCCATCGGC AGTGTGGAC TGGGGAAGGT CCTCATAGAC ATCCTTGAG GGTATGGCGC
 CCGCGACCGA ATCGACCGC GCGGTAGCCG TCACAACCTG ACCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGCG

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L
 3841. GGGCGTGGCG GGAGCTCTTG TGGCATTCAA GATCATGAGC GGTGAGGTCC CCTCCACGGA GGACCTGGTC AATCTACTGC
 CCCGCACCGC CCTCGAGAAC ACCGTAAGTT CTAGTACTCG CCACTCCAGG GGAGGTGCCT CCTGGACCAG TTAGATGACG

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E
 3921. CCGCCATCCT CTCGCCCCGGA GCCCTCGTAG TCGGCGTGGT CTGTGCAGCA ATACTGCGCC GGCACGTTGG CCCGGGCGAG
 GGGCGTAGGA GAGCGGGCCT CGGGAGCATC AGCCGCACCA GACACGTCGT TATGACGCGG CCGTGCAACC GGGCCCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E
 4001. GGGGCGTGC AGTGGATGAA CCGGCTGATA GCCTTCGCCT CCCGGGGGAA CCATGTTTCC CCCACGCACT ACGTGCCGGA
 CCCCCTCAG TCACCTACTT GGCCGACTAT CGGAAGCGGA GGGCCCCCTT GGTACAAAGG GGGTGCGTGA TGCACGGCCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W
 4081. GAGCGATGCA GCTGCCCCGCG TCACTGCCAT ACTCAGCAGC CTCACTGTAA CCCAGTCCT GAGGCGACTG CACCAGTGGA
 CTCGTACGT CGACGGGCGC AGTGACGGTA TGAGTCGTCG GAGTGACATT GGGTCGAGGA CTCCGTGAC GTGGTCACCT

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D
 4161. TAAGCTCGGA GTGTACCACT CCATGCTCCG GTTCTGGCT AAGGGACATC TGGGACTGGA TATGCGAGGT GTTGAGCGAC
 ATTCGAGCCT CACATGGTGA GGTACGAGGC CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACTCGCTG

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G
 BamHI
 ~~~~~  
 4241. TTTAAGACCT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCGCG GGTATAAGGG  
 AAATTCTGGA CCGATTTTCG ATTCGAGTAC GGTGTGCGAC GACCCTAGGG GAAACACAGG ACGGTCGCGC CCATATTCCC

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T  
 4321. GGTCTGGCGA GGGGACGGCA TCATGCACAC TCGCTGCCAC TGTGGAGCTG AGATCACTGG ACATGTCAAA AACGGGACGA  
 CCAGACCGCT CCCCTGCCGT AGTACGTGTG AGCGACGGTG ACACCTCGAC TCTAGTGACC TGTACAGTTT TTGCCCTGCT

+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C  
 4401. TGAGGATCGT CGGTCCTAGG ACCTGCAGGA ACATGTGGAG TGGGACCTTC CCCATTAATG CCTACACCAC GGGCCCCGTG  
 ACTCTAGCA GCCAGGATCC TGGACGTCCT TGTACACCTC ACCCTGGAAG GGGTAATTAC GGATGTGGTG CCCGGGGACA

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G  
 4481. ACCCCCCCTT CTGCGCCGAA CTACACGTTT GCGCTATGGA GGGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGGG  
 TGGGGGGAAG GACGCGGCTT GATGTGCAAG CGCGATACCT CCCACAGACG TCTCCTTATG CACCTCTATT CCGTCCACCC

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T  
 4561. GGAATCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCCCT GCCAGTCCC ATCGCCCGAA TTTTTCACAG  
 CCTGAAGGTG ATGCACTGCC CATACTGATG ACTGTTAGAA TTTACGGGCA CGGTCCAGGG TAGCGGGCTT AAAAAGTGTC

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G  
 4641. AATTGGACGG GGTGCGCCTA CATAGGTTTG CGCCCCCTG CAAGCCCTTG CTGCGGGAGG AGGTATCATT CAGAGTAGGA  
 TTAACCTGCC CCACGCGGAT GTATCCAAAC GCGGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCCT

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D  
 4721. CTCCACGAAT ACCCGGTAGG GTCGCAATTA CCTTGCGAGC CCGAACCAGG CGTGCCCGTG TTGACGTCCA TGCTCACTGA  
 GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGCTCG GGCTTGGCCT GCACCGGCAC AACTGCAGGT ACGAGTGACT

+2 P S H I T A E A A G R R L A R G S P P S V A S S S A  
 4801. TCCCTCCCAT ATAACAGCAG AGGCGGGCCG GCGAAGGTTG GCGAGGGGAT CACCCCCCTC TGTGGCCAGC TCCTCGGCTA  
 AGGGAGGGTA TATTGTCGTC TCCGCCGGCC CGCTTCCAAC CGTCCCCTA GTGGGGGGAG ACACCGGTG AGGAGCCGAT

## FIGURE 5 - Page 5

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N  
 4881 GCCAGCTATC CGCTCCATCT CTCAAGGCAA CTTGCACCGC TAACCATGAC TCCCCTGATG CTGAGCTCAT AGAGGCCAAC  
 CCGTTCGATAG GCGAGGTAGA GAGTTCGGTT GAACGTGGCG ATTGGTACTG AGGGGACTAC GACTCGAGTA TCTCCGGTTG

---

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D  
 4961 CTCCTATGGA GGCAGGAGAT GGGCGGCAAC ATCACCAGGG TTGAGTCAGA AAACAAAGTG GTGATTCTGG ACTCCTTCGA  
 GAGGATACCT CCGTCCTCTA CCCGCCGTTG TAGTGGTCCC AACTCAGTCT TTTGTTTCAC CACTAAGACC TGAGGAAGCT

---

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q  
 5041 TCCGCTTGTG GCGGAGGAGG ACGAGCGGGA GATCTCCGTA CCCGCAGAAA TCCTGCGGAA GTCTCGGAGA TTCGCCAGG  
 AGGCGAACAC CGCTCCTCC TGCTCGCCCT CTAGAGGCAT GGGCGTCTT AGGACGCCTT CAGAGCCTCT AAGCGGGTCC

---

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V  
 5121 CCCTGCCCCGTTGGGCGCGG CCGGACTATA ACCCCCCGCT AGTGGAGACG TGGAAAAAGC CCGACTACGA ACCACCTGTG  
 GGGACGGGCA AACCCGCGCC GGCCTGATAT TGGGGGGCGA TCACCTCTGC ACCTTTTTCG GGCTGATGCT TGGTGGACAC

---

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E  
 5201 GTCCATGGCT GCCCGCTTCC ACCTCCAAAG TCCCCTCCTG TGCTCCGCC TCGGAAGAAG CGGACGGTGG TCCTCACTGA  
 CAGGTACCGA CGGGCGAAGG TGGAGGTTT AGGGGAGGAC ACGGAGGCGG AGCCTTCTTC GCCTGCCACC AGGAGTGACT

---

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D  
 5281 ATCAACCCTA TCTACTGCCT TGGCCGAGCT CGCCACCAGA AGCTTTGGCA GCTCCTCAAC TTCCGGCATT ACGGGCGACA  
 TAGTTGGGAT AGATGACGGA ACCGGCTCGA GCGGTGGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTAA TGCCCGCTGT

---

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P  
 5361 ATACGACAAC ATCCTCTGAG CCCGCCCTT CTGGCTGCCC CCCCAGACTCC GACGCTGAGT CCTATTCTC CATGCCCCCC  
 TATGCTGTTG TAGGAGACTC GGGCGGGGAA GACCGACGGG GGGGCTGAGG CTGCGACTCA GGATAAGGAG GTACGGGGGG

---

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V  
 BamHI  
 ~~~~~  
 5441 CTGGAGGGGG AGCCTGGGGA TCCGATCTT AGCGACGGGT CATGGTCAAC GGTCAGTAGT GAGGCCAACG CGGAGGATGT
 GACCTCCCCC TCGGACCCCT AGGCCTAGAA TCGCTGCCCA GTACCAAGTTG CCAGTCATCA CTCCGGTTGC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I
 5521 CGTGTGCTGC TCAATGTCTT ACTCTTGGAC AGGCGCACTC GTCACCCCGT GCGCCGCGGA AGAACAGAAA CTGCCCATCA
 GCACACGACG AGTTACAGAA TGAGAACCTG TCCGCGTGAG CAGTGGGGCA CGCGGCGCCT TCTTGCTTTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K
 5601 ATGCACTAAG CAACTCGTTG CTACGTCACC ACAATTTGGT GTATTCCACC ACCTCAGCA GTGCTTGCCA AAGGCAGAAG
 TACGTGATTC GTTGAGCAAC GATGCAGTGG TGTTAAACCA CATAAGGTGG TGGAGTGCGT CACGAACGGT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A A S K
 5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA
 TTTCAGTGTA AACTGTCTGA CGTTCAAGAC CTGTCGGTAA TGGTCTGCA TGAGTTCCTC CAATTCGTC GCCGCAAGTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y
 5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGACGC CTGACGCCCC CACACTCAGC CAAATCCAAG TTTGGTTATG
 TCACTTCCGA TTGAACGATA GGCATCTCCT TCGAACGTCG GACTGCGGGG GTGTGAGTCG GTTTAGGTTT AAACCAATAC

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N
 5841 GGGCAAAAGA CGTCCGTTGC CATGCCAGAA AGGCCGTAAC CCACATCAAC TCCGTGTGGA AAGACCTTCT GGAAGACAAT
 CCCGTTTTCT GCAGGCAACG GTACGGTCTT TCCGGCATTG GGTGTAGTTG AGGCACACCT TTCTGGAAGA CTTTCTGTTA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A
 5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCCTGAG AAGGGGGGTC GTAAGCCAGC
 CATTGTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCAG CATTGCGTCG

FIGURE 5 - Page 6

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P
 6001 TCGTCTCATC GTGTTCCCCG ATCTGGGCGT GCGCGTGTGC GAAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCCT
 AGCAGAGTAG CACAAGGGGC TAGACCCGCA CGCGCACACG CTTTCTACC GAAACATGCT GCACCAATGT TTCGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S
 EcoRI
 ~~~~~

6081 TGGCCGTGAT GGAAGCTCC TACGGATTCC AATACTCACC AGGACAGCGG GTTGAATTCC TCGTGCAAGC GTGGAAGTCC  
 ACCGGCACTA CCCTTCGAGG ATGCCTAAGG TTATGAGTGG TCCTGTCGCC CAACTTAAGG AGCACGTTCC CACCTTCAGG

---

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E  
 6161 AAGAAAACCC CAATGGGGTT CTCGTATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGCGACATCC GTACGGAGGA  
 TTCTTTTGGG GTTACCCCAA GAGCATACTA TGGGCGACGA AACTGAGGTG TCAGTGACTC TCGCTGTAGG CATGCCTCCT

---

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G  
 6241 GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG  
 CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTCGGGCG CACCGGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCC

---

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G  
 6321 GCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGCTATCG CAGGTGCCGC GCGAGCGGCG TACTGACAAC TAGCTGTGGT  
 CGGGAGAATG GTTAAGTTCC CCCCTCTTGA CGCGATAGC GTCCACGGCG CGCTCGCCGC ATGACTGTTG ATCGACACCA

---

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G  
 6401 AACACCCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCGAGG GCTCCAGGAC TGCACCATGC TCGTGTGTGG  
 TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCTG ACGTGGTACG AGCACACACC

---

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M  
 6481 CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGGTCCAG GAGGACGCGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA  
 GCTGCTGAAT CAGCAATAGA CACTTTCGCG CCCCAGGTC CTCCTGCGCC GCTCGGACTC TCGGAAGTGC CTCCGATACT

---

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V  
 6561 CCAGGTACTC CGCCCCCCT GGGGACCCCG CACAACCAGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTG  
 GGTCCATGAG GCGGGGGGGA CCCCTGGGGG GTGTTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC

---

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W  
 6641 TCAGTCGCCC ACGACGGCGC TGGAAAGAGG GTCTACTACC TCACCGTGA CCCTACAACC CCCCTCGCGA GAGCTGCGTG  
 AGTCAGCGGG TGCTGCCGCG ACCTTTCTCC CAGATGATGG AGTGGGCACT GGGATGTTGG GGGGAGCGCT CTCGACGCAC

---

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M  
 6721 GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTTGCCCC CACACTGTGG GCGAGGATGA  
 CCTCTGTCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAAACGGGG GTGTGACACC CGCTCCTACT

---

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A  
 6801 TACTGATGAC CCATTTCTTT AGCGTCCTTA TAGCCAGGGA CCAGCTTGAA CAGGCCCTCG ATTGCGAGAT CTACGGGGCC  
 ATGACTACTG GGTAAAGAAA TCGCAGGAAT ATCGGTCCTT GGTCGAACTT GTCCGGGAGC TAACGCTCTA GATGCCCCGG

---

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y  
 6881 TGCTACTCCA TAGAACCACT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA  
 ACGATGAGGT ATCTTGGTGA CCTAGATGGA GGTTAGTAAG TTTCTGAGGT ACCGGAGTCG CGTAAAAGTG AGGTGTCAAT

---

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R  
 6961 CTCTCCAGGT GAAATCAATA GGGTGGCCGC ATGCCTCAGA AAACCTGGGG TACCGCCCTT GCGAGCTTGG AGACACGGGG  
 GAGAGGTCCA CTTTAGTTAT CCCACGGGCG TACGGAGTCT TTTGAACCCC ATGGCGGGAA CGCTCGAACC TCTGTGGCCC

---

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V  
 7041 CCCGGAGCGT CCGCGCTAGG CTTCTGGCCA GAGGAGGCAG GGCTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA  
 GGGCCTCGCA GCGCGATCC GAAGACCGGT CTCCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT

8401 TGATCTTTTC TACGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG  
ACTAGAAAAG ATGCCCCAGA CTGGAGTCA CCTTGCTTTT GAGTGCATTT CCTTAAAACC AGTACTCTAA TAGTTTTTCC

[illegible]

## FIGURE 5 - Page 8

8481 ATCTTCACCT AGATCCTTTT AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG  
TAGAAGTGGA TCTAGGAAAA TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA CCAGACTGTC

---

8561 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTT GTTCATCCAT AGTTGCCTGA CTCCCCGTCG  
AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC

---

8641 TG TAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT  
ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACCGGG GTCACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA

---

8721 CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA  
GGTCTAAATA GTCGTTATTT GGTGCGTCGG CCTTCCCGGC TCGCGTCTTC ACCAGGACGT TGAAATAGGC GGAGGTAGGT

---

8801 GTCTATTAAT TGTGCGGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG  
CAGATAATTA ACAACGGCCC TTCGATCTCA TTCATCAAGC GGTCAATTAT CAAACGCGTT GCAACAACGG TAACGATGTC

---

8881 GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC  
CGTAGCACCA CAGTGCGAGC AGCAAACCAT ACCGAAGTAA GTCGAGGCCA AGGGTTGCTA GTTCCGCTCA ATGTACTAGG

---

8961 CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTCGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTATCACT  
GGGTACAACA CGTTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATTC AACCGGCGTC ACAATAGTGA

---

9041 CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA  
GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAAG AACTGACCA CTCATGAGTT

---

9121 CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT  
GGTTCAGTAA GACTCTTATC ACATACGCCG CTGGCTCAAC GAGAACGGGC CGCAGTTATG CCCTATTATG GCGCGGTGTA

---

9201 AGCAGAACTT TAAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTGAGATC  
TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTTGCAAGAA GCCCCGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

---

9281 CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA  
GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTTT

---

9361 CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTCAATAT  
GTCCTTCCGT TTTACGGCGT TTTTTCCTT ATTCCCGCTG TGCCTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

---

9441 TATTGAAGCA TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC AAATAGGGGT  
ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAATC TACATAAATC TTTTATTTG TTTATCCCCA

---

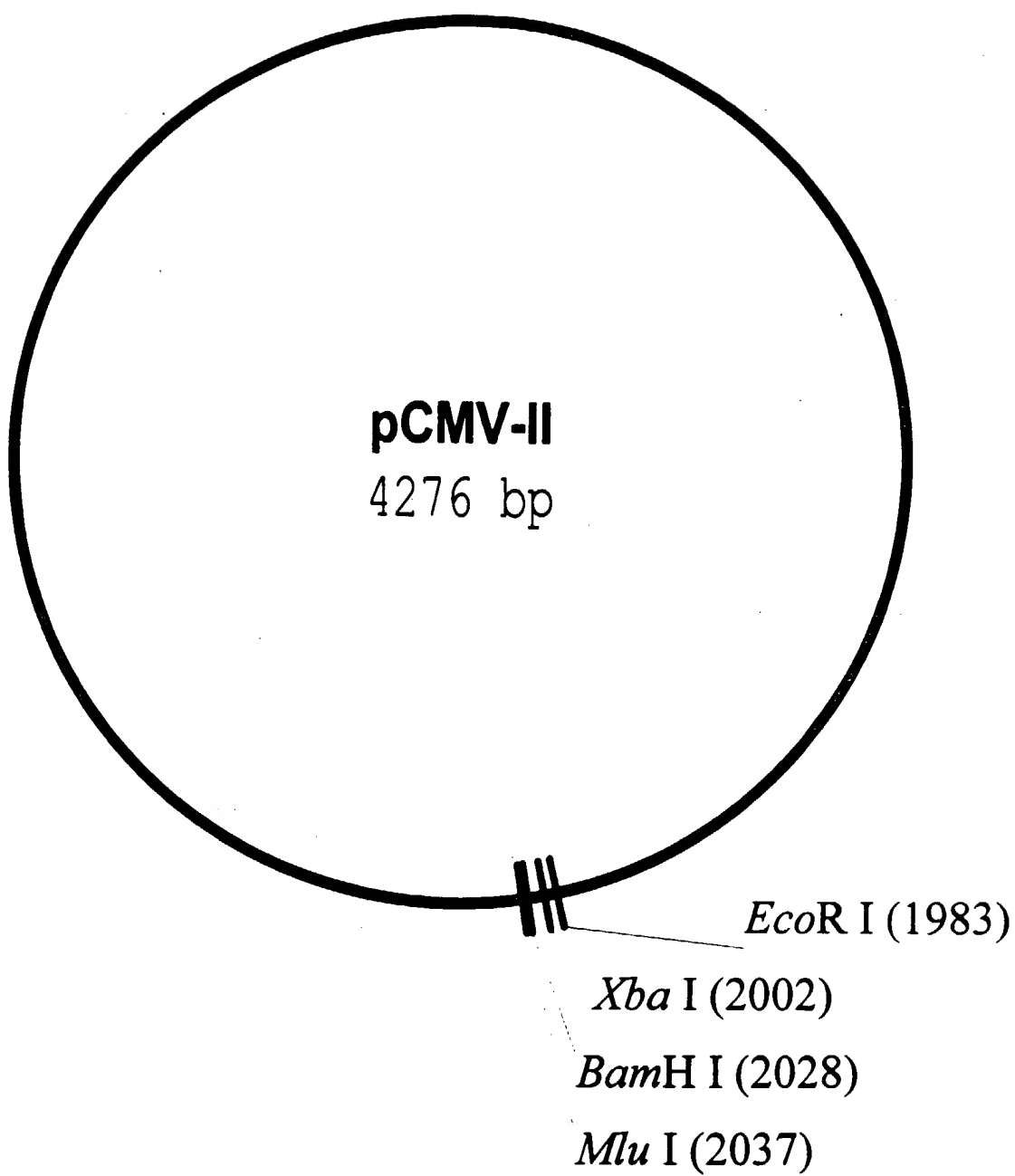
9521 TCCGCGCACA TTTCCCCGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGC  
AGGCGCGTGT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTTGTAAT AATAGTACTG TAATTGGATA TTTTATCCG

---

9601 GTATCACGAG GCCCTTTCGT C  
CATAGTGCTC CGGAAAAGCA G

0923476-4200

FIGURE 6





## FIGURE 7 - Page 1

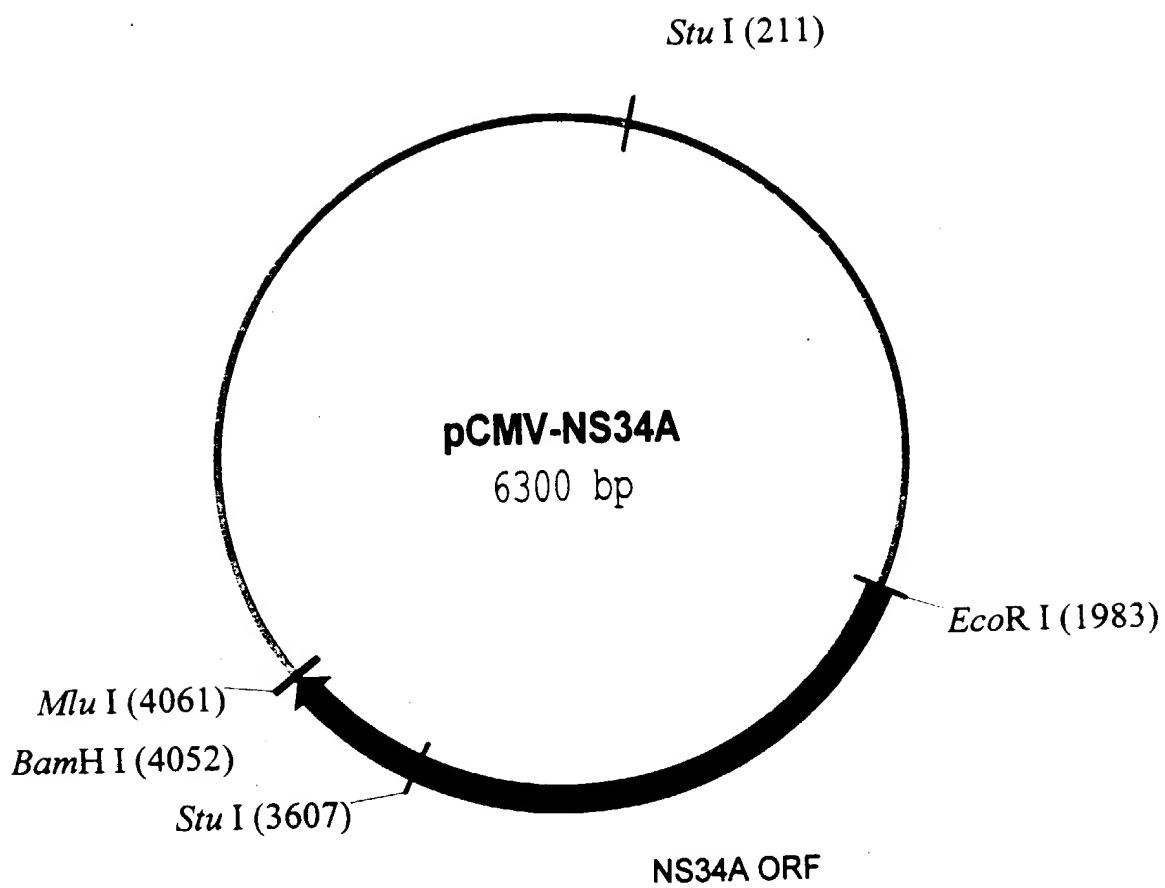
|      |            |            |            |            |            |            |             |            |
|------|------------|------------|------------|------------|------------|------------|-------------|------------|
| 1    | TCGCGCGTTT | CGGTGATGAC | GGTGAAAACC | TCTGACACAT | GCAGCTCCCG | GAGACGGTCA | CAGCTTGTCT  | GTAAGCGGAT |
|      | AGCGCGCAAA | GCCACTACTG | CCACTTTTGG | AGACTGTGTA | CGTCGAGGGC | CTCTGCCAGT | GTCGAACAGA  | CATTGCGCTA |
| 81   | GCCGGGAGCA | GACAAGCCCC | TCAGGGCGCG | TCAGCGGGTG | TTGGCGGGTG | TCGGGGCTGG | CTTAACATATG | CGGCATCAGA |
|      | CGGCCCTCGT | CTGTTCCGGC | AGTCCCGCGC | AGTCGCCAC  | AACCGCCCAC | AGCCCCGACC | GAATTGATAC  | GCCGTAGTCT |
| 161  | GCAGATTGTA | CTGAGAGTGC | ACCATATGAA | GCTTTTTGCA | AAAGCCTAGG | CCTCCAAAAA | AGCCTCCTCA  | CTACTTCTGG |
|      | CGTCTAACAT | GACTCTCACG | TGGTATACTT | CGAAAAACGT | TTTCGGATCC | GGAGGTTTTT | TCGGAGGAGT  | GATGAAGACC |
| 241  | AATAGCTCAG | AGGCCGAGGC | GGCCTCGGCC | TCTGCATAAA | TAAAAAAAAT | TAGTCAGCCA | TGGGGCGGAG  | AATGGGCGGA |
|      | TTATCGAGTC | TCCGGCTCCG | CCGGAGCCGG | AGACGTATTT | ATTTTTTTTA | ATCAGTCGGT | ACCCCGCCTC  | TTACCCGCCT |
| 321  | ACTGGGCGGG | GAGGGAATTA | TTGGCTATTG | GCCATTGCAT | ACGTTGTATC | TATATCATAA | TATGTACATT  | TATATTGGCT |
|      | TGACCCGCCC | CTCCCTTAAT | AACCGATAAC | CGGTAACGTA | TGCAACATAG | ATATAGTATT | ATACATGTAA  | ATATAACCGA |
| 401  | CATGTCCAAT | ATGACCGCCA | TGTTGACATT | GATTATTGAC | TAGTTATTAA | TAGTAATCAA | TTACGGGGTC  | ATTAGTTCAT |
|      | GTACAGGTTA | TACTGGCGGT | ACAACGTAA  | CTAATAACTG | ATCAATAATT | ATCATTAGTT | AATGCCCCAG  | TAATCAAGTA |
| 481  | AGCCCATATA | TGGAGTTCCG | CGTTACATAA | CTTACGGTAA | ATGGCCCGCC | TGGCTGACCG | CCCAACGACC  | CCCGCCCAT  |
|      | TCGGGTATAT | ACCTCAAGGC | GCAATGTATT | GAATGCCATT | TACCGGGCGG | ACCGACTGGC | GGGTGCTGG   | GGGCGGGTAA |
| 561  | GACGTCAATA | ATGACGTATG | TTCCCATAGT | AACGCCAATA | GGGACTTTCC | ATTGACGTCA | ATGGGTGGAG  | TATTTACGGT |
|      | CTGCAGTTAT | TACTGCATAC | AAGGGTATCA | TTGCGGTTAT | CCCTGAAAGG | TAAGTGCAGT | TACCCACCTC  | ATAAATGCCA |
| 641  | AAACTGCCCC | CTTGGCAGTA | CATCAAGTGT | ATCATATGCC | AAGTCCGCCC | CCTATTGACG | TCAATGACGG  | TAAATGGCCC |
|      | TTTGACGGGT | GAACCGTCAT | GTAGTTTACA | TAGTATACGG | TTCAGGCGGG | GGATAACTGC | AGTTACTGCC  | ATTTACCGGG |
| 721  | GCCTGGCATT | ATGCCAGTA  | CATGACCTTA | CGGGACTTTC | CTACTTGGCA | GTACATCTAC | GTATTAGTCA  | TCGCTATTAC |
|      | CGGACCGTAA | TACGGGTCAT | GTAAGTGAAT | GCCCTGAAAG | GATGAACCGT | CATGTAGATG | CATAATCAGT  | AGCGATAATG |
| 801  | CATGGTGATG | CGGTTTTGGC | AGTACACCAA | TGGGCGTGGA | TAGCGGTTTG | ACTCACGGGG | ATTTCCAAGT  | CTCCACCCCA |
|      | GTACCACTAC | GCCAAAACCG | TCATGTGGTT | ACCCGCACCT | ATCGCCAAAC | TGAGTGCCCC | TAAAGGTTCA  | GAGGTGGGGT |
| 881  | TTGACGTCAA | TGGGAGTTTG | TTTTGGCACC | AAAATCAACG | GGACTTTCCA | AAATGTCGTA | ATAACCCCGC  | CCCGTTGACG |
|      | AAGTGCAGTT | ACCCTCAAAC | AAAACCGTGG | TTTAGTTGTC | CCTGAAAGGT | TTTACAGCAT | TATTGGGGCG  | GGGCAACTGC |
| 961  | CAAATGGGCG | GTAGGCGTGT | ACGGTGGGAG | GTCTATATAA | GCAGAGCTCG | TTTAGTGAAC | CGTCAGATCG  | CCTGGAGACG |
|      | GTTTACCCGC | CATCCGCACA | TGCCACCCTC | CAGATATATT | CGTCTCGAGC | AAATCACTTG | GCAGTCTAGC  | GGACCTCTGC |
| 1041 | CCATCCACGC | TGTTTTGACC | TCCATAGAAG | ACACCGGGAC | CGATCCAGCC | TCCGCGGCGG | GGAACGGTGC  | ATTGGAACGC |
|      | GGTAGGTGCG | ACAAAACCTG | AGGTATCTTC | TGTGGCCCTG | GCTAGGTCCG | AGGCGCCGGC | CCTTGCCACG  | TAACCTTGCG |
| 1121 | GGATTCCCCG | TGCCAAGAGT | GACGTAAGTA | CCGCTATAG  | ACTCTATAGG | CACACCCCTT | TGGCTCTTAT  | GCATGCTATA |
|      | CCTAAGGGGC | ACGGTTCTCA | CTGCATTCTA | GGCGGATATC | TGAGATATCC | GTGTGGGGAA | ACCGAGAATA  | CGTACGATAT |
| 1201 | CTGTTTTTGG | CTTGGGGCCT | ATACACCCCC | GCTCCTTATG | CTATAGGTGA | TGGTATAGCT | TAGCCTATAG  | GTGTGGGTTA |
|      | GACAAAACCC | GAACCCCGGA | TATGTGGGGG | CGAGGAATAC | GATATCCACT | ACCATATCGA | ATCGGATATC  | CACACCCAAT |
| 1281 | TTGACCATTA | TTGACCACTC | CCCTATTGGT | GACGATACTT | TCCATTACTA | ATCCATAACA | TGGCTCTTTG  | CCACAACAT  |
|      | AAGTGGTAAT | AAGTGGTGAG | GGGATAACCA | CTGCTATGAA | AGGTAATGAT | TAGGTATTGT | ACCGAGAAAC  | GGTGTGATA  |
| 1361 | CTCTATTGGC | TATATGCCAA | TACTCTGTCC | TTCAGAGACT | GACACGGACT | CTGTATTTTT | ACAGGATGGG  | GTCCATTTAT |
|      | GAGATAACCG | ATATACGGTT | ATGAGACAGG | AAGTCTCTGA | CTGTGCCTGA | GACATAAAAA | TGTCCTACCC  | CAGGTAAATA |
| 1441 | TATTTACAAA | TTCACATATA | CAACAACGCC | GTCCCCCGTG | CCCGCAGTTT | TTATTAAACA | TAGCGTGGGA  | TCTCCGACAT |
|      | ATAAATGTTT | AAGTGTATAT | GTTGTTGCGG | CAGGGGGCAC | GGGCGTCAAA | AATAATTTGT | ATCGCACCTC  | AGAGGCTGTA |

## FIGURE 7 - Page 2

|                                                                                                                          |            |            |            |            |              |             |             |            |            |             |            |            |            |            |             |            |
|--------------------------------------------------------------------------------------------------------------------------|------------|------------|------------|------------|--------------|-------------|-------------|------------|------------|-------------|------------|------------|------------|------------|-------------|------------|
| 1521                                                                                                                     | CTCGGGTACG | TGTTCCGGAC | ATGGGCTCTT | CTCCGGTAGC | GGCGGAGCTT   | CCACATCCGA  | GCCCTGGTCC  | CATCCGTCCA | GAGCCCATGC | ACAAGGCCTG  | TACCCGAGAA | GAGGCCATCG | CCGCCTCGAA | GGTGTAGGCT | CGGGACCAGG  | GTAGGCAGGT |
| 1601                                                                                                                     | GCGGCTCATG | GTCGCTCGGC | AGCTCCTTGC | TCCTAACAGT | GGAGGCCAGA   | CTTAGGCACA  | GCACAATGCC  | CACCACCACC | CGCCGAGTAC | CAGCGAGCCG  | TCGAGGAACG | AGGATTGTCA | CCTCCGGTCT | GAATCCGTGT | CGTGTTACGG  | GTGGTGGTGG |
| 1681                                                                                                                     | AGTGTGCCGC | ACAAGGCCGT | GGCGGTAGGG | TATGTGTCTG | AAAATGAGCT   | CGGAGATTGG  | GCTCGCACCT  | GGACGCAGAT | TCACACGGCG | TGTTCCGGCA  | CCGCCATCCC | ATACACAGAC | TTTTACTCGA | GCCTCTAACC | CGAGCGTGGA  | CCTGCGTCTA |
| 1761                                                                                                                     | GGAAGACTTA | AGGCAGCGGC | AGAAGAAGAT | GCAGGCAGCT | GAGTTGTTGT   | ATTCTGATAA  | GAGTCAGAGG  | TAATCCCGT  | CCTTCTGAAT | TCCGTGCGCG  | TCTTCTTCTA | CGTCCGTCGA | CTCAACAACA | TAAGACTATT | CTCAGTCTCC  | ATTGAGGGCA |
| 1841                                                                                                                     | TGCGGTGCTG | TTAACGGTGG | AGGGCAGTGT | AGTCTGAGCA | GTA CTCTGTTG | CTGCCGCGCG  | CGCCACCAGA  | CATAATAGCT | ACGCCACGAC | AATTGCCACC  | TCCCGTCACA | TCAGACTCGT | CATGAGCAAC | GACGGCGCGC | GCGGTGGTCT  | GTATTATCGA |
| EcoRI<br>~~~~~                                                                                                           |            |            |            |            |              |             |             |            |            |             |            |            |            |            |             |            |
| 1921                                                                                                                     | GACAGACTAA | CAGACTGTTC | CTTTCCATGG | GTCTTTTCTG | CAGTCACCGT   | CGTCGACCTA  | AGAATTTCAGA | CTCGAGCAAG | CTGTCTGATT | GTCTGACAAAG | GAAAGGTACC | CAGAAAAGAC | GTCAGTGGCA | GCAGCTGGAT | TCTTAAGTCT  | GAGCTCGTTC |
| XbaI                      BamHI                      MluI<br>~~~~~                      ~~~~~                      ~~~~~ |            |            |            |            |              |             |             |            |            |             |            |            |            |            |             |            |
| 2001                                                                                                                     | TCTAGAAAGG | CGCGCCAAGA | TATCAAGGAT | CCACTACGCG | TTAGAGCTCG   | CTGATCAGCC  | TCGACTGTGC  | CTTCTAGTTG | AGATCTTTCC | GCGCGGTTCT  | ATAGTTCCTA | GGTGATGCGC | AATCTCGAGC | GACTAGTCGG | AGCTGACACG  | GAAGATCAAC |
| 2081                                                                                                                     | CCAGCCATCT | GTTGTTTGCC | CCTCCCCCGT | GCCTTCCTTG | ACCCTGGAAG   | GTGCCACTCC  | CACTGTCCTT  | TCCTAATAAA | GGTCGGTAGA | CAACAAACGG  | GGAGGGGGCA | CGGAAGGAAC | TGGGACCTTC | CACGGTGAGG | GTGACAGGAA  | AGGATTATTT |
| 2161                                                                                                                     | ATGAGGAAAT | TGCATCGCAT | TGTCTGAGTA | GGTGTCAATC | TATTCTGGGG   | GGTGGGGTGG  | GGCAGGACAG  | CAAGGGGGAG | TACTCCTTTA | ACGTAGCGTA  | ACAGACTCAT | CCACAGTAAG | ATAAGACCCC | CCACCCACC  | CCGTCTGTGC  | GTTCCCCCTC |
| 2241                                                                                                                     | GATTGGGAAG | ACAATAGCAG | GCATGCTGGG | GAGCTCTTCC | GCTTCTCTCG   | TCAGTACTC   | GCTGCGCTCG  | GTCGTTCCGC | CTAACCTTTC | TGTTATCGTC  | CGTACGACCC | CTCGAGAAGG | CGAAGGAGCG | AGTACTGAG  | CGACGCGAGC  | CAGCAAGCCG |
| 2321                                                                                                                     | TGCGGCGAGC | GGTATCAGCT | CACTCAAAGG | CGGTAATACG | GTTATCCACA   | GAATCAGGGG  | ATAACGCAGG  | AAAGAACATG | ACGCCGCTCG | CCATAGTCGA  | GTGAGTTTCC | GCCATTATGC | CAATAGGTGT | CTTAGTCCCC | TATTGCGTCC  | TTTCTTGTAC |
| 2401                                                                                                                     | TGAGCAAAAG | GCCAGCAAAA | GGCCAGGAAC | CGTAAAAAGG | CCGCGTTGCT   | GGCGTTTTTC  | CATAGGCTCC  | GCCCCCTGA  | ACTCGTTTTC | CGGTCGTTTT  | CCGGTCCTTG | GCATTTTTCC | GGCGCAACGA | CCGCAAAAAG | GTATCCGAGG  | CGGGGGGACT |
| 2481                                                                                                                     | CGAGCATCAC | AAAAATCGAC | GCTCAAGTCA | GAGGTGGCGA | AACCCGACAG   | GA CTATAAAG | ATACCAGGCG  | TTTCCCCCTG | GCTCGTAGTG | TTTTTAGCTG  | CGAGTTCAGT | CTCCACCGCT | TTGGGCTGTC | CTGATATTTT | TATGGTCCGC  | AAAGGGGGAC |
| 2561                                                                                                                     | GAAGCTCCCT | CGTGCGCTCT | CCTGTTCCGA | CCCTGCCGCT | TACCGGATAC   | CTGTCCGCCT  | TTCTCCCTTC  | GGGAAGCGTG | CTTCGAGGGA | GCACGCGAGA  | GGACAAGGCT | GGGACGGCGA | ATGGCCTATG | GACAGGCGGA | AAGAGGGAAG  | CCCTTCGCAC |
| 2641                                                                                                                     | GCGCTTTCTC | AATGCTCAGC | CTGTAGGTAT | CTCAGTTCGG | TGTAGGTCGT   | TCGCTCCAAG  | CTGGGCTGTG  | TGCACGAACC | CGCGAAAGAG | TTACGAGTGC  | GACATCCATA | GAGTCAAGCC | ACATCCAGCA | AGCGAGGTTC | GACCCGACAC  | ACGTGCTTGG |
| 2721                                                                                                                     | CCCCGTTTCT | CCCGACCGCT | GCGCCTTATC | CGGTAAGTAT | CGTCTTGAGT   | CCAACCCGGT  | AAGACACGAC  | TTATCGCCAC | GGGGCAAGTC | GGGCTGGCGA  | CGCGGAATAG | GCCATTGATA | GCAGAACTCA | GGTTGGGCCA | TTCTGTGCTG  | AATAGCGGTG |
| 2801                                                                                                                     | TGGCAGCAGC | CACTGGTAAC | AGGATTAGCA | GAGCGAGGTA | TGTAGGCGGT   | GCTACAGAGT  | TCTTGAAGTG  | GTGGCCTAAC | ACCGTCGTCG | GTGACCATTG  | TCCTAATCGT | CTCGCTCCAT | ACATCCGCCA | CGATGTCTCA | AGAACTTCAC  | CACCGGATTG |
| 2881                                                                                                                     | TACGGCTACA | CTAGAAGGAC | AGTATTTGGT | ATCTGCGCTC | TGCTGAAGCC   | AGTTACCTTC  | GGAAAAAGAG  | TTGGTAGCTC | ATGCCGATGT | GATCTTCCTG  | TCATAAACCA | TAGACGCGAG | ACGACTTCGG | TCAATGGAAG | CCTTTTTTCTC | AACCATCGAG |

|      |                          |                           |                           |                           |                            |                          |                           |                           |
|------|--------------------------|---------------------------|---------------------------|---------------------------|----------------------------|--------------------------|---------------------------|---------------------------|
| 2961 | TTGATCCGGC<br>AACTAGGCCG | AAACAAACCA<br>TTTGTGTTGGT | CCGCTGGTAG<br>GGCGACCATC  | CGGTGGTTTT<br>GCCACCAAAA  | TTTGTGTTGCA<br>AAACAAACGT  | AGCAGCAGAT<br>TCGTCGTCTA | TACGCGCAGA<br>ATGCGCGTCT  | AAAAAAGGAT<br>TTTTTTCCTA  |
| 3041 | CTCAAGAAGA<br>GAGTTCCTCT | TCCTTTGATC<br>AGGAAACTAG  | TTTTCTACGG<br>AAAAGATGCC  | GGTCTGACGC<br>CCAGACTGCG  | TCAGTGGAAC<br>AGTCACCTTG   | GAAAACACAC<br>CTTTTGAGTG | GTTAAGGGAT<br>CAATTCCCTA  | TTTGGTTCATG<br>AAACCAGTAC |
| 3121 | AGATTATCAA<br>TCTAATAGTT | AAAGGATCTT<br>TTTCCTAGAA  | CACCTAGATC<br>GTGGATCTAG  | CTTTTAAATT<br>GAAAATTTAA  | AAAAATGAAG<br>TTTTTACTTC   | TTTTAAATCA<br>AAAATTTAGT | ATCTAAAGTA<br>TAGATTTTCAT | TATATGAGTA<br>ATATACTCAT  |
| 3201 | AACTTGGTCT<br>TTGAACCAGA | GACAGTTACC<br>CTGTCAATGG  | AATGCTTAAT<br>TTACGAATTA  | CAGTGAGGCA<br>GTCACTCCGT  | CCTATCTCAG<br>GGATAGAGTC   | CGATCTGTCT<br>GCTAGACAGA | ATTTTCGTTCA<br>TAAAGCAAGT | TCCATAGTTG<br>AGGTATCAAC  |
| 3281 | CCTGACTCCC<br>GGACTGAGGG | CGTCGTGTAG<br>GCAGCACATC  | ATAACTACGA<br>TATTGATGCT  | TACGGGAGGG<br>ATGCCCTCCC  | CTTACCATCT<br>GAATGGTAGA   | GGCCCCAGTG<br>CCGGGGTCAC | CTGCAATGAT<br>GACGTTACTA  | ACCGCGAGAC<br>TGGCGCTCTG  |
| 3361 | CCACGCTCAC<br>GGTGCGAGTG | CGGCTCCAGA<br>GCCGAGGTCT  | TTTATCAGCA<br>AAATAGTCGT  | ATAAACCAGC<br>TATTTGGTCG  | CAGCCGGAAG<br>GTCGGCCTTC   | GGCCGAGCGC<br>CCGGCTCGCG | AGAAGTGGTC<br>TCTTCACCAG  | CTGCAACTTT<br>GACGTTGAAA  |
| 3441 | ATCCGCCTCC<br>TAGGCGGAGG | ATCCAGTCTA<br>TAGGTCAGAT  | TTAATTGTTG<br>AATTAACAAC  | CCGGGAAGCT<br>GGCCCTTCGA  | AGAGTAAGTA<br>TCTCATTTCAT  | GTTCCGCAGT<br>CAAGCGGTCA | TAATAGTTTG<br>ATTATCAAAC  | CGCAACGTTG<br>GCGTTGCAAC  |
| 3521 | TTGCCATTGC<br>AACGGTAACG | TACAGGCATC<br>ATGTCCGTAG  | GTGGTGTAC<br>CACCACAGTG   | GCTCGTCGTT<br>CGAGCAGCAA  | TGGTATGGCT<br>ACCATACCGA   | TCATTTCAGT<br>AGTAAGTCGA | CCGGTTCCCA<br>GGCCAAGGGT  | ACGATCAAGG<br>TGCTAGTTCC  |
| 3601 | CGAGTTACAT<br>GCTCAATGTA | GATCCCCCAT<br>CTAGGGGGTA  | GTTGTGCAAA<br>CAACACGTTT  | AAAGCGGTTA<br>TTTCGCCAAT  | GCTCCTTCGG<br>CGAGGAAGCC   | TCCTCCGATC<br>AGGAGGCTAG | GTTGTCAGAA<br>CAACAGTCTT  | GTAAGTTGGC<br>CATTCAACCG  |
| 3681 | CGCAGTGTTA<br>GCGTCACAAT | TCACTCATGG<br>AGTGAGTACC  | TTATGGCAGC<br>AATACCGTCG  | ACTGCATAAT<br>TGACGTATTA  | TCTCTTACTG<br>AGAGAATGAC   | TCATGCCATC<br>AGTACGGTAG | CGTAAGATGC<br>GCATTCTACG  | TTTTCTGTGA<br>AAAAGACACT  |
| 3761 | CTGGTGAGTA<br>GACCACTCAT | CTCAACCAAG<br>GAGTTGGTTC  | TCATTCTGAG<br>AGTAAGACTC  | AATAGTGTAT<br>TTATCACATA  | GCGGCGACCG<br>CGCCGCTGGC   | AGTTGCTCTT<br>TCAACGAGAA | GCCC GGCGTC<br>CGGGCCGCAG | AATACGGGAT<br>TTATGCCCTA  |
| 3841 | AATACCGCGC<br>TTATGGCGCG | CACATAGCAG<br>GTGTATCGTC  | AACTTTAAAA<br>TTGAAATTTT  | GTGCTCATCA<br>CACGAGTAGT  | TTGGA AAAACG<br>AACCTTTTGC | TTCTTCGGGG<br>AAGAAGCCCC | CGAAAACCTCT<br>GCTTTTGAGA | CAAGGATCTT<br>GTTCTAGAA   |
| 3921 | ACCGCTGTTG<br>TGGCGACAAC | AGATCCAGTT<br>TCTAGGTCAA  | CGATGTAACC<br>GCTACATTGG  | CACTCGTGCA<br>GTGAGCACGT  | CCCAACTGAT<br>GGGTTGACTA   | CTTCAGCATC<br>GAAGTCGTAG | TTTTACTTTT<br>AAAATGAAAG  | ACCAGCGTTT<br>TGGTCGCAAA  |
| 4001 | CTGGGTGAGC<br>GACCCACTCG | AAAAACAGGA<br>TTTTTGTCTT  | AGGCAAAATG<br>TCCGTTTTTAC | CCGCAAAAAA<br>GGCGTTTTTTT | GGAATAAAGG<br>CCCTTATTCC   | GCGACACGGA<br>CGCTGTGCCT | AATGTTGAAT<br>TTACAACCTA  | ACTCATACTC<br>TGAGTATGAG  |
| 4081 | TTCCTTTTTT<br>AAGGAAAAAG | AATATTATTG<br>TTATAATAAC  | AAGCATTTTAT<br>TTCGTAAATA | CAGGGTTATT<br>GTCCCAATAA  | GTCTCATGAG<br>CAGAGTACTC   | CGGATACATA<br>GCCTATGTAT | TTTGAATGTA<br>AAACTTACAT  | TTTAGAAAAA<br>AAATCTTTTT  |
| 4161 | TAAACAAATA<br>ATTTGTTTAT | GGGGTTCCGC<br>CCCCAAGGCG  | GCACATTTCC<br>CGTGTA AAGG | CCGAAAAGTG<br>GGCTTTTTCAC | CCACCTGACG<br>GGTGACTGC    | TCTAAGAAAC<br>AGATTCTTTG | CATTATTATC<br>GTAATAATAG  | ATGACATTAA<br>TACTGTAATT  |
| 4241 | CCTATAAAAA<br>GGATATTTTT | TAGGCGTATC<br>ATCCGCATAG  | ACGAGGCCCT<br>TGCTCCGGGA  | TTCGTC<br>AAGCAG          |                            |                          |                           |                           |

FIGURE 8



## FIGURE 9 - Page 1

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG  
AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

---

51 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG  
CTCTGCCAGT GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTGCGGG

---

101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG  
AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

---

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGTGA  
GCCGTAGTCT CGTCTAACAT GACTCTCAG TGGTATACTT CGAAAAACGT

---

StuI  
~~~~~

201 AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG
TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

251 AGGCCGAGGC GGCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA
TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

301 TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG
ACCCCGCCTC TTACCCGCCT TGACCCGCCC CTCCCTTAAT AACCGATAAC

351 GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401 CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

451 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCGG
ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

501 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

551 CCCGCCCAT TACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA
GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

601 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAAGTGGCCA
CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

651 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG
GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

701 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA
AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

751 CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

801 CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC

851 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG
TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC

FIGURE 9 - Page 2

901 TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC
AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG

951 CCCGTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA
GGGCAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT

1001 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC
CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG

1051 TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG
ACAAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGCG AGGCGCCGGC

1101 GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA
CCTTGCCACG TAACCTTGCG CTAAGGGGC ACGGTCTCA CTGCATTCA

1151 CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCG GCTCCTTATG CTATAGGTGA
GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT

1251 TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC
ACCATATCGA ATCGGATATC CACACCCAAT AACTGGTAAT AACTGGTGAG

1301 CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG
GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC

1351 CCACAACATAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT
GGTGTTGATA GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA

1401 GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT TATTTACAAA
CTGTGCCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA ATAAATGTTT

1451 TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCGAGTTT TTATTAAACA
AAGTGTATAT GTTGTTCGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT

1501 TAGCGTGGA TCTCCGACAT CTCGGGTACG TGTTCCGGAC ATGGGCTCTT
ATCGCACCTT AGAGGCTGTA GAGCCCATGC ACAAGGCCTG TACCCGAGAA

1551 CTCCGGTAGC GCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

1651 CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT
GAATCCGTGT CGTGTACGG GTGGTGGTGG TCACACGGCG GTTCCGGCA

1701 GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT
CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA

1751 GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT
CCTGCGTCTA CCTTCTGAAT TCCGTGCGCG TCTTCTCTA CGTCCGTGCA

1801 GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TCGGGTGCTG
CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC

FIGURE 9 - Page 3

1851 TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG
AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

1901 CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG
GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

+2 M A P
EcoRI
~~~~~

1951 GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCGCCCA  
CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGCGGGT

---

+2 I T A Y A Q Q T R G L L G C I I T

2001 TCACGGCGTA CGCCCAGCAG ACAAGGGGCC TCCTAGGGTG CATAATCACC  
AGTGCCGCAT GCGGGTCGTC TGTTCGCCG AGGATCCAC GTATTAGTGG

---

+2 S L T G R D K N Q V E G E V Q I V

2051 AGCCTAACTG GCCGGGACAA AAACCAAGTG GAGGGTGAGG TCCAGATTGT  
TCGGATTGAC CGGCCCTGTT TTTGGTTCAC CTCCCACTCC AGGTCTAACA

---

+2 S T A A Q T F L A T C I N G V C

2101 GTCAACTGCT GCCCAAACCT TCCTGGCAAC GTGCATCAAT GGGGTGTGCT  
CAGTTGACGA CGGGTTTGGA AGGACCGTTG CACGTAGTTA CCCCACAGA

---

+2 W T V Y H G A G T R T I A S P K G

2151 GGA CTGTCTA CCACGGGGCC GGAACGAGGA CCATCGCGTC ACCCAAGGGT  
CCTGACAGAT GGTGCCCCCG CTTTGCTCCT GGTAGCGCAG TGGTTTCCCA

---

+2 P V I Q M Y T N V D Q D L V G W P

2201 CCTGTCATCC AGATGTATAC CAATGTAGAC CAAGACCTTG TGGGCTGGCC  
GGACAGTAGG TCTACATATG GTTACATCTG GTTCTGGAAC ACCCGACCGG

---

+2 A S Q G T R S L T P C T C G S S

2251 CGCTTCGCAA GGTACCCGCT CATTGACACC CTGCACTTGC GGCTCCTCGG  
GCGAAGCGTT CCATGGGCGA GTAACGTGG GACGTGAACG CCGAGGAGCC

---

+2 D L Y L V T R H A D V I P V R R R

2301 ACCTTTACCT GGTACGAGG CACGCCGATG TCATTCCCGT GCGCCGGCGG  
TGGAAATGGA CCAGTGCTCC GTGCGGCTAC AGTAAGGGCA CGCGGCCGCC

---

+2 G D S R G S L L S P R P I S Y L K

2351 GGTGATAGCA GGGGCAGCCT GCTGTCGCCC CGGCCCATTT CCTACTTGAA  
CCACTATCGT CCCCCTCGGA CGACAGCGGG GCCGGGTAAA GGATGAAC TT

---

+2 G S S G G P L L C P A G H A V G

2401 AGGCTCCTCG GGGGGTCCGC TGTGTGCCC CGCGGGGCAC GCCGTGGGCA  
TCCGAGGAGC CCCCAGGCG ACAACACGGG GCGCCCCGTG CGGCACCCGT

---

+2 I F R A A V C T R G V A K A V D F

2451 TATTTAGGGC CGCGGTGTGC ACCCGTGGAG TGGCTAAGGC GGTGGACTTT  
ATAAATCCCC GCGCCACAG TGGGCACCTC ACCGATTCCG CCACCTGAAA

---

+2 I P V E N L E T T M R S P V F T D

2501 ATCCCTGTGG AGAACCTAGA GACAACCATG AGGTCCCCGG TGTTACGGA  
TAGGGACACC TCTTGGATCT CTGTTGGTAC TCCAGGGGCC ACAAGTGCCT

PCMV-NS34A  
FIGURE 9 - Page 4

+2 N S S P P V V P Q S F Q V A H L  
2551 TAACTCCTCT CCACCAGTAG TGCCCCAGAG CTTCCAGGTG GCTCACCTCC  
ATTGAGGAGA GGTGGTCATC ACGGGGTCTC GAAGGTCCAC CGAGTGGAGG

---

+2 H A P T G S G K S T K V P A A Y A  
2601 ATGCTCCAC AGGCAGCGGC AAAAGCACCA AGGTCCCGGC TGCATATGCA  
TACGAGGGTG TCCGTCGCCG TTTTCGTGGT TCCAGGGCCG ACGTATACGT

---

+2 A Q G Y K V L V L N P S V A A T L  
2651 GCTCAGGGCT ATAAGGTGCT AGTACTCAAC CCCTCTGTTG CTGCAACACT  
CGAGTCCCGA TATTCCACGA TCATGAGTTG GGGAGACAAC GACGTTGTGA

---

+2 G F G A Y M S K A H G I D P N I  
2701 GGGCTTTGGT GCTTACATGT CCAAGGCTCA TGGGATCGAT CCTAACATCA  
CCCGAAACCA CGAATGTACA GGTTCGAGT ACCCTAGCTA GGATTGTAGT

---

+2 R T G V R T I T T G S P I T Y S T  
2751 GGACCGGGT GAGAACAATT ACCACTGGCA GCCCCATCAC GTACTCCACC  
CCTGGCCCCA CTCTTGTTAA TGGTGACCGT CGGGGTAGTG CATGAGGTGG

---

+2 Y G K F L A D G G C S G G A Y D I  
2801 TACGGCAAGT TCCTTGCCGA CGGCGGGTGC TCGGGGGGCG CTTATGACAT  
ATGCCGTTCA AGGAACGGCT GCCGCCACG AGCCCCCGC GAATACTGTA

---

+2 I I C D E C H S T D A T S I L G  
2851 AATAATTTGT GACGAGTGCC ACTCCACGGA TGCCACATCC ATCTTGGGCA  
TTATTAAACA CTGCTACGG TGAGGTGCCT ACGGTGTAGG TAGAACCCGT

---

+2 I G T V L D Q A E T A G A R L V V  
2901 TTGGCACTGT CCTTGACCAA GCAGAGACTG CGGGGGCGAG ACTGGTTGTG  
AACCCTGACA GGAAGTGGT CGTCTCTGAC GCCCCCGCTC TGACCAACAC

---

+2 L A T A T P P G S V T V P H P N I  
2951 CTCGCCACCG CCACCCCTCC GGGCTCCGTC ACTGTGCCCC ATCCCAACAT  
GAGCGGTGGC GGTGGGAGG CCCGAGGAG TGACACGGGG TAGGGTTGTA

---

+2 E E V A L S T T G E I P F Y G K  
3001 CGAGGAGGTT GCTCTGTCCA CCACCGGAGA GATCCCTTTT TACGGCAAGG  
GCTCCTCAA CGAGACAGGT GGTGGCCTCT CTAGGGA AAA ATGCCGTTCC

---

+2 A I P L E V I K G G R H L I F C H  
3051 CTATCCCCCT CGAAGTAATC AAGGGGGGGA GACATCTCAT CTTCTGTCAT  
GATAGGGGGA GCTTCATTAG TTCCCCCCT CTGTAGAGTA GAAGACAGTA

---

+2 S K K K C D E L A A K L V A L G I  
3101 TCAAAGAAGA AGTGCGACGA ACTCGCCGCA AAGCTGGTGC CATTGGGCAT  
AGTTTCTTCT TCACGCTGCT TGAGCGGCGT TTCGACCAGC GTAACCCGTA

---

+2 N A V A Y Y R G L D V S V I P T  
3151 CAATGCCGTG GCCTACTACC GCGGTCTTGA CGTGTCCGTC ATCCCGACCA  
GTTACGGCAC CGGATGATGG CGCCAGAACT GCACAGGCAG TAGGGCTGGT

---

+2 S G D V V V V A T D A L M T G Y T  
3201 GCGGCGATGT TGTCGTCGTG GCAACCGATG CCCTCATGAC CGGCTATACC  
CGCCGCTACA ACAGCAGCAC CGTTGGCTAC GGGAGTACTG GCCGATATGG

0923479-11200



## FIGURE 9 - Page 5

+2 G D F D S V I D C N T C V T Q T V  
3251 GCGACTTCG ACTCGGTGAT AGACTGCAAT ACGTGTGTCA CCCAGACAGT  
CCGCTGAAGC TGAGCCACTA TCTGACGTTA TGCACACAGT GGGTCTGTCA

+2 D F S L D P T F T I E T I T L P  
3301 CGATTTTCAGC CTTGACCCTA CTTTACCCTA TGAGACAATC ACGTCCCCC  
GCTAAAGTCG GAACTGGGAT GGAAGTGGTA ACTCTGTAG TGCGAGGGGG

+2 Q D A V S R T Q R R G R T G R G K  
3351 AAGATGCTGT CTCCCGCACT CAACGTCGGG GCAGGACTGG CAGGGGGAAG  
TTCTACGACA GAGGGCGTGA GTTGACGCCC CGTCCTGACC GTCCCCCTC

+2 P G I Y R F V A P G E R P S G M F  
3401 CCAGGCATCT ACAGATTTGT GGCACCGGGG GAGCGCCCCT CCGGCATGTT  
GGTCCGTAGA TGTCTAAACA CCGTGGCCCC CTCGCGGGGA GGCCGTACAA

+2 D S S V L C E C Y D A G C A W Y  
3451 CGACTCGTCC GTCCTCTGTG AGTGCTATGA CGCAGGCTGT GCTTGGTATG  
GCTGAGCAGG CAGGAGACAC TCACGATACT GCGTCCGACA CGAACCATAC

+2 E L T P A E T T V R L R A Y M N T  
3501 AGCTCACGCC CGCCGAGACT ACAGTTAGGC TACGAGCGTA CATGAACACC  
TCGAGTGCGG GCGGCTCTGA TGTCAATCCG ATGCTCGCAT GTACTTGTGG

+2 P G L P V C Q D H L E F W E G V F  
3551 CCGGGGCTTC CCGTGTGCCA GGACCATCTT GAATTTTGGG AGGGCGTCTT  
GGCCCCGAAG GGCACACGGT CCTGGTAGAA CTAAACCC TCCCGCAGAA

+2 T G L T H I D A H F L S Q T K Q  
StuI  
~~~~~  
3601 TACAGGCCTC ACTCATATAG ATGCCCACTT TCTATCCCAG ACAAAGCAGA
ATGTCCGGAG TGAGTATATC TACGGGTGAA AGATAGGGTC TGTTTCGTCT

+2 S G E N L P Y L V A Y Q A T V C A
3651 GTGGGGAGAA CCTTCCTTAC CTGGTAGCGT ACCAAGCCAC CGTGTGCGCT
CACCCCTCTT GGAAGGAATG GACCATCGCA TGGTTCGGTG GCACACGCA

+2 R A Q A P P P S W D Q M W K C L I
3701 AGGGCTCAAG CCCCTCCCC ATCGTGGGAC CAGATGTGGA AGTGTGTTGAT
TCCCGAGTTC GGGGAGGGGG TAGCACCCTG GTCTACACCT TCACAACTA

+2 R L K P T L H G P T P L L Y R L
3751 TCGCCTCAAG CCCACCCTCC ATGGGCCAAC ACCCCTGCTA TACAGACTGG
AGCGGAGTTC GGGTGGGAGG TACCCGGTTG TGGGGACGAT ATGTCTGACC

+2 G A V Q N E I T L T H P V T K Y I
3801 GCGCTGTTCA GAATGAAATC ACCCTGACGC ACCCAGTCAC CAAATACATC
CGCGACAAGT CTTACTTTAG TGGGACTGCG TGGGTCACTG GTTTATGTAG

+2 M T C M S A D L E V V T S T W V L
3851 ATGACATGCA TGTCGGCCGA CCTGGAGGTC GTCACGAGCA CCTGGGTGCT
TACTGTACGT ACAGCCGGCT GGACCTCCAG CAGTGCTCGT GGACCCACGA

+2 V G G V L A A L A A Y C L S T G
3901 CGTTGGCGGC GTCCTGGCTG CTTTGGCCGC GTATTGCCTG TCAACAGGCT
GCAACCGCCG CAGGACCGAC GAAACCGGCG CATAACGGAC AGTTGTCCGA

FIGURE 9 - Page 6

4801 CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT
GCCATTCTGT GCTGAATAGC GGTGACCGTC GTCGGTGACC ATTGTCCTAA

[illegible]

FIGURE 9 - Page 7

4851 AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC
TCGTCTCGCT CCATACATCC GCCACGATGT CTCAAGAACT TCACCACCGG

4901 TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA
ATTGATGCCG ATGTGATCTT CCTGTCATAA ACCATAGACG CGAGACGACT

4951 AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA
TCGGTCAATG GAAGCCTTTT TCTCAACCAT CGAGAACTAG GCCGTTTGTT

5001 ACCACCGCTG GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG
TGGTGGCGAC CATCGCCACC AAAAAAACAA ACGTTCGTCG TCTAATGCGC

5051 CAGAAAAAAA GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG
GTCTTTTTTT CCTAGAGTTC TTCTAGGAAA CTAGAAAAGA TGCCCCAGAC

5101 ACGCTCAGTG GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA
TGGAGTCAC CTTGCTTTTG AGTGCAATTC CCTAAAACCA GTACTCTAAT

5151 TCAAAAAGGA TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTTAA
AGTTTTTCCT AGAAGTGGAT CTAGGAAAAT TTAATTTTTA CTTCAAAATT

5201 ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT
TAGTTAGATT TCATATATAC TCATTTGAAC CAGACTGTCA ATGGTTACGA

5251 TAATCAGTGA GGCACCTATC TCAGCGATCT GTCTATTTTCG TTCATCCATA
ATTAGTCACT CCGTGGATAG AGTCGCTAGA CAGATAAAGC AAGTAGGTAT

5301 GTTGCCTGAC TCCCCGTCGT GTAGATAACT ACGATACGGG AGGGCTTACC
CAACGGACTG AGGGGCAGCA CATCTATTGA TGCTATGCCC TCCCGAATGG

5351 ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC
TAGACCGGGG TCACGACGTT ACTATGGCGC TCTGGGTGCG AGTGGCCGAG

5401 CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT
GTCTAAATAG TCGTTATTTG GTCGGTCGGC CTTCCCGGCT CGCGTCTTCA

5451 GGTCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA
CCAGGACGTT GAAATAGGCG GAGGTAGGTC AGATAATTAA CAACGGCCCT

5501 AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTTGCCA
TCGATCTCAT TCATCAAGCG GTCAATTATC AAACGCGTTG CAACAACGGT

5551 TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTTGGTAT GGCTTCATT
AACGATGTCC GTAGACCAC AGTGCGAGCA GCAAACCATA CCGAAGTAAG

5601 AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG
TCGAGGCCAA GGGTTGCTAG TTCCGCTCAA TGTACTAGG GGTACAACAC

5651 CAAAAAAGCG GTTAGCTCCT TCGGTCCTCC GATCGTTGTC AGAAGTAAGT
GTTTTTTCGC CAATCGAGGA AGCCAGGAGG CTAGCAACAG TCTTCATTCA

5701 TGGCCGCACT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT
ACCGGCGTCA CAATAGTGAG TACCAATACC GTCGTGACGT ATTAAGAGAA

5751 ACTGTCATGC CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC
TGACAGTACG GTAGGCATTC TACGAAAAGA CACTGACCAC TCATGAGTTG

FIGURE 9 - Page 8

5801 CAAGTCATTG TGAGAATAGT GTATGCGGCG ACCGAGTTGC TCTTGCCCCG
GTTCAAGTAAG ACTCTTATCA CATACGCCGC TGGCTCAACG AGAACGGGCC

5851 CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC
GCAGTTATGC CCTATTATGG CGCGGTGTAT CGTCTTGAAA TTTTCACGAG

5901 ATCATTGGAA AACGTTCTTC GGGGCGAAAA CTCTCAAGGA TCTTACCGCT
TAGTAACCTT TTGCAAGAAG CCCCCTTTT GAGAGTTCCT AGAATGGCGA

5951 GTTGAGATCC AGTTCGATGT AACCCTACTCG TGCACCCAAC TGATCTTCAG
CAACTCTAGG TCAAGCTACA TTGGGTGAGC ACGTGGGTG ACTAGAAGTC

6001 CATCTTTTAC TTTCAACAGC GTTCTGCGGT GAGCAAAAAC AGGAAGGCAA
GTAGAAAATG AAAGTGGTCG CAAAGACCCA CTCGTTTTTG TCCTTCCGTT

6051 AATGCCGCAA AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT
TTACGGCGTT TTTTCCCTTA TTCCCGCTGT GCCTTTACAA CTTATGAGTA

6101 ACTCTTCCTT TTTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA
TGAGAAGGAA AAAGTTATAA TAACTTCGTA AATAGTCCCA ATAACAGAGT

6151 TGAGCGGATA CATATTTGAA TGTATTTAGA AAAATAAACA AATAGGGGTT
ACTCGCCTAT GTATAAACTT ACATAAATCT TTTTATTTGT TTATCCCCAA

6201 CCGCGCACAT TTCCCCGAAA AGTGCCACCT GACGTCTAAG AAACCATTAT
GGCGCGTGTA AAGGGGCTTT TCACGGTGGA CTGCAGATTC TTTGGTAATA

6251 TATCATGACA TTAACCTATA AAAATAGGCG TATCACGAGG CCCTTTCGTC
ATAGTACTGT AATTGGATAT TTTTATCCGC ATAGTGCTCC GGGAAAGCAG

Diagram 1

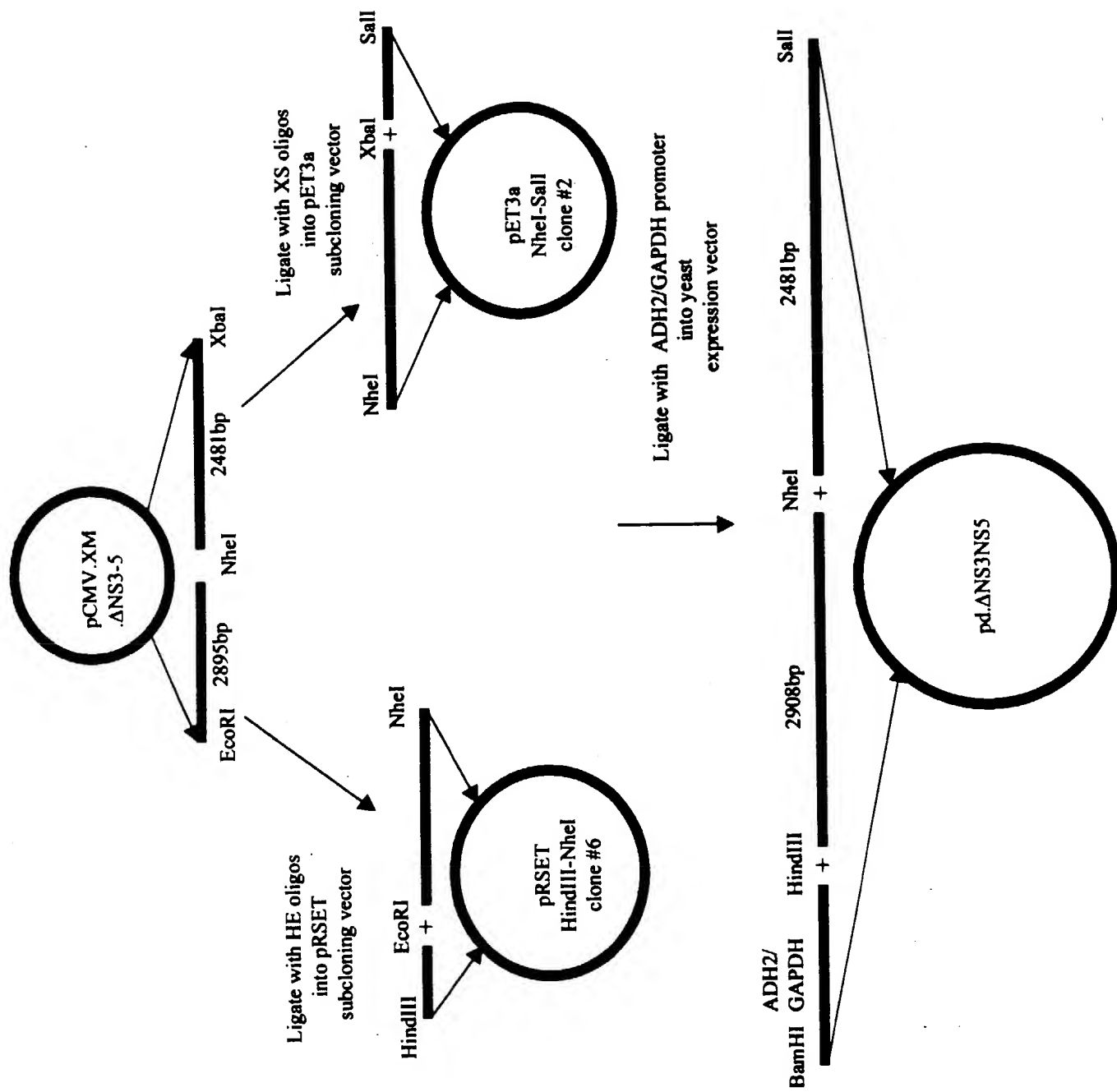


FIGURE 11 - Page 1

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuVal
 2 AGCTTACAAAACAAATTCACCATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTA
 TCGAATGTTTTGTTTAAGTGGTACCGACGTATACGTCGAGTCCCGATATTCCACGATCAT
 ^ ^ ^ ^
 1 HIND3, 21 NCOI, 30 NDEI, 58 SCAI,

LeuAsnProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGly
 62 CTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGG
 GAGTTGGGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCC

IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyr
 122 ATCGATCCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTAC
 TAGCTAGGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATG
 ^
 122 CLAI,

SerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIle
 182 TCCACCTACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATA
 AGGTGGATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTAT

IleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeu
 242 ATTTGTGACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTT
 TAAACACTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAA

AspGlnAlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGly
 302 GACCAAGCAGAGACTGCGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGC
 CTGGTTCGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGAGGCCCC
 ^
 309 ALWN1,

SerValThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIle
 362 TCCGTCACTGTGCCCCATCCCAACATCGAGGAGGTGCTCTGTCCACCACCGGAGAGATC
 AGGCAGTGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAG

ProPheTyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePhe
 422 CCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTC
 GGAAAAATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAG

CysHisSerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsn
 482 TGTCATTCAAAGAAGAAGTGCGACGAACCTGCGCGAAAGCTGGTCGCATTGGGCATCAAT
 ACAGTAAGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTA

AlaValAlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValVal
 542 GCCGTGGCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCGGCGATGTTGTC
 CGGCACCGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAG
 ^ ^
 556 SAC2, 566 DRD1,

ValValAlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAsp
 602 GTCGTGGCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGAC
 CAGCACCGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTG
 ^
 621 BSPH1,

CysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGlu

FIGURE 11 - Page 2

662 TGCAATACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAG
ACGTTATGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAACGGGATGGAAGTGGTAACTC

722 ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArg
ACAATCACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGG
TGTTAGTGCAGAGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCC

782 GlyLysProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAsp
GGGAAGCCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCCCTCCGGCATGTTTCGAC
CCCTTCGGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCGGTACAAGCTG
^ ^

822 BGLI, 839 DRD1,

842 SerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAla
TCGTCCGTCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCC
AGCAGGCAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCAGGGCGG
^

887 SACI,

902 GluThrThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAsp
GAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGAC
CTCTGATGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTTG
^

937 SMAI XMAI,

962 HisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeu
CATCTTGAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCCACTTTCTA
GTAGAACTTAAAACCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGAT
^

991 STUI,

1022 SerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrVal
TCCCAGACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTG
AGGGTCTGTTTTCGTCTCACCCCTCTTGAAGGAATGGACCATCGCATGGTTCGGTGGCAC
^

1075 DRA3,

1082 CysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArg
TGCGCTAGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTCGC
ACGCGATCCCGAGTTCGGGGAGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCG

1142 LeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsn
CTCAAGCCCCACCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAGAAT
GAGTTCGGGTGGGAGGTACCCGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTA
^

1156 NCOI,

1202 GluIleThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeu
GAAATCACCCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTG
CTTTAGTGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGAC
^ ^ ^ ^ ^

1236 BSPH1, 1240 DRD1, 1243 AVA3, 1251 EAG1 XMA3, 1256 DRD1,

1262 GluValValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyr
GAGGTCGTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTAT
CTCCAGCAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATA

Downloaded from www.sciencedirect.com

FIGURE 11 - Page 3

1322 CysLeuSerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAla
 TGCCTGTCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCA
 ACGGACAGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGT
 1375 NAEI,
 1382 IleIleProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGln
 ATCATACCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGTCTCTCAG
 TAGTATGGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTC
 1391 DRD1,
 1442 HisLeuProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeu
 CACTTACCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTC
 GTGAATGGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAG
 1502 GlyLeuLeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsn
 GGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGTGTCCAGACCAAC
 CCGGAGGACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTG
 1508 PSTI, 1513 TTH3I,
 1562 TrpGlnLysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGln
 TGGCAAAAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAAGTTCATCAGTGGGATACAA
 ACCGTTTTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTT
 1571 XHOI, 1592 NDEI,
 1622 TyrLeuAlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPhe
 TACTTGGCGGGCTTGTC AACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTT
 ATGAACCGCCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACACCGAAAA
 1649 BSTE2,
 1682 ThrAlaAlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGly
 ACAGCTGCTGTCAACGCTGCTGCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTT
 TGTGACGACAGTGGTTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCC
 1683 ALWN1 PVU2,
 1742 GlyTrpValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGly
 GGGTGGGTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGC
 CCCACCCACCGACGGGTCGAGCGGGCGGGGCCACGGCGATGACGGAAACACCCGCGACCG
 1800 ESP1,
 1802 LeuAlaGlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAla
 TTAGCTGGCGCCGCGCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCA
 AATCGACCGCGCGGTAGCCGTACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGT
 1808 KAS1 NARI,
 1862 GlyTyrGlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluVal
 GGGTATGGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGACGGGTGAGGTC
 CCCATACCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAG

ProSerThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuVal
1922 CCCTCCACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTA
GGGAGGTGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCAT
1934 TTH3I,
ValGlyValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaVal
1982 GTCGGCGTGGTCTGTGCAGCAATACTGCGCCGGCAGCTTGGCCCGGGCAGGGGGCAGTG
CAGCCGCACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCAC
2010 NAEI, 2023 SMAI XMAI,
GlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHis
2042 CAGTGGATGAACCGGCTGATAGCCTTCGCCCTCCCGGGGAACCATGTTTCCCCACGCAC
GTCACCTACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTG
2073 SMAI XMAI, 2099 DRA3,
TyrValProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrVal
2102 TACGTGCCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTA
ATGCACGGCCTCTCGCTACGTTCGACGGGCGCAGTGACGGTATGAGTCGTGCGAGTGACAT
2121 PVU2,
ThrGlnLeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSer
2162 ACCCAGCTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTAACCACTCCATGCTCC
TGGGTTCGAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGG
2165 ALWN1, 2170 MST2,
GlySerTrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThr
2222 GGTTCCTGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACC
CCAAGGACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGG
2226 ECON1,
TrpLeuLysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArg
2282 TGGCTAAAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCTGCCAGCGC
ACCGATTTTCGATTTCGAGTACGGTGTGCGACGGACCCTAGGGGAAACACAGGACGGTCGCG
2291 ESP1, 2306 PVU2, 2316 BAMHI,
GlyTyrLysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAla
2342 GGGTATAAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCT
CCCATATTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGA
2402 GluIleThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArg
GAGATCACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGG
CTCTAGTGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCC
2431 BSAB1, 2447 AVR2, 2454 SSE83871, 2455 PSTI,
AsnMetTrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeu
2462 AACATGTGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCTGTACCCCCCTT
TTGTACACCTCACCTGGAAGGGGTAATTACGGATGTGGTGGCCGGGGACATGGGGGGAA

2486 ASE1, 2503 APAI,

2522 ProAlaProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIle
CCTGCGCCGAACACTACAGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATA
GGACGCGGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTAT

2559 PSTI,

2582 ArgGlnValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysPro
AGGCAGGTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCG
TCCGTCCACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGC

2600 DRA3,

2642 CysGlnValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPhe
TGCCAGGTCCCATCGCCGAATTTTTCACAGAATTGGACGGGGTGCCTACATAGGTTT
ACGGTCCAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGGGATGTATCCAAA

2702 AlaProProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGlu
GCGCCCCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTGAGAGTAGGACTCCACGAA
CGCGGGGGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTT

2762 TyrProValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSer
TACCCGGTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCC
ATGGGCCATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACCTGCAGG

2763 HGIE2, 2815 AAT2,

2822 MetLeuThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGly
ATGCTCACTGATCCCTCCCATATAACAGCAGAGGCGCGGGCGAAGGTTGGCGAGGGGA
TACGAGTGAAGTAGGGAGGTATATTGTCGTCTCCGCCGCGCCGCTTCCAACCGCTCCCTT

2856 EAG1 XMA3,

2882 SerProProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAla
TCACCCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCA
AGTGGGGGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGT

2895 BALI, 2909 NHEI,

2942 ThrCysThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrp
ACTTGACCCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGG
TGAACGTGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACC

2972 ESP1, 2975 SACI,

3002 ArgGlnGluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeu
AGGCAGGAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGATTCTG
TCCGTCTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGAC

3062 AspSerPheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGlu
GACTCCTTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGAGAA
CTGAGGAAGCTAGGCGAACACCGCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTT

3102 BGL2,

FIGURE 11 - Page 6

3122 IleLeuArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyr
 ATCCTGCGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCGTTTGGGCGCGGCCGGACTAT
 TAGGACGCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGCCTGATA
 3149 ALWN1, 3170 EAG1 XMA3,
 AsnProProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGly
 3182 AACCCCCCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGC
 TTGGGGGGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCG
 3223 HGIE2, 3235 NCOI,
 CysProLeuProProProLysSerProProValProProProArgLysLysArgThrVal
 3242 TGCCCGCTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTG
 ACGGGCGAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCAC
 ValLeuThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGly
 3302 GTCCTCACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGC
 CAGGAGTGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAACCG
 3338 SACI, 3352 HIND3,
 SerSerSerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaPro
 3362 AGCTCCTCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCT
 TCGAGGAGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGA
 SerGlyCysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGly
 3422 TCTGGCTGCCCCCGGACTCCGACGCTGAGTCTTATCCTCCATGCCCCCCTGGAGGGG
 AGACCGACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCC
 3443 EAM11051,
 GluProGlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsn
 3482 GAGCCTGGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAAC
 CTCGGACCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTG
 3490 BAMHI, 3491 BSAB1, 3493 BSPE1,
 AlaGluAspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrPro
 3542 GCGGAGGATGTCGTGTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCG
 CGCCTCCTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGC
 3595 DRA3,
 CysAlaAlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHis
 3602 TGCGCCGCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCAC
 ACGCGGCGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTGTTGAGCAACGATGCAGTG
 3606 SAC2, 3617 ALWN1, 3661 PFLM1,
 HisAsnLeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThr
 3662 CACAATTTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAAGTCA
 GTGTTAAACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTCAGTGT
 3687 DRA3,
 PheAspArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAla

3722 TTTGACAGACTGCAAGTTCCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCA
AAACTGTCTGACGTTCAAGACCTGTCTGGTAATGGTCTCTGCATGAGTTCCTCCAATTCGT

3782 AlaAlaSerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrPro
GCGGCGTCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCC
CGCCGCAGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTGCGACTGCGGG
3822 HIND3,

3842 ProHisSerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArg
CCACACTCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGA
GGTGTGAGTCGGTTTAGGTTCAAACCAATACCCCGTTTCTGCAGGCAACGGTACGGTCT
3881.AAT2, 3896 BGLI,

3902 LysAlaValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrPro
AAGGCCGTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCA
TTCCGGCATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGT

3962 IleAspThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGly
ATAGACACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGT
TATCTGTGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCA

4022 ArgLysProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMet
CGTAAGCCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATG
GCATTGCGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTAC

4082 AlaLeuTyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPhe
GCTTTGTACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCT
CGAAACATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAG

4142 GlnTyrSerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThr
CAATACTCACCAGGACAGCGGGTTGAATTCTCGTGCAAGCGTGGAAGTCCAAGAAAACC
GTTATGAGTGGTCTGTGCCCCAACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGG
4166 ECORI,

4202 ProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIle
CCAATGGGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATC
GGTTACCCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCTAGTACTCTCGCTGTAG
4235 DRD1, 4242 ALWN1,

4262 ArgThrGluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIle
CGTACGGAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGGTGGCCATC
GCATGCCTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTTCGGGCGCACCGGTAG
4307 BGLI, 4314 BALI,

4322 LysSerLeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsn
AAGTCCCTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAC
TTCAGGGAGTGGCTCTCCGAAATACAACCCCGGAGAAATGGTTAAGTTCCCCCTCTTG
4351 APAI,

4382 CysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeu
TGCGGCTATCGCAGGTGCCGCGCAGCGGCGTACTGACAACTAGCTGTGGTAACACCTCT

[illegible]

FIGURE 11 - Page 9

5042 ValProProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAla
GTACCGCCCTTGCGAGCTTGGAGACACCGGGCCGGAGCGTCCGCGCTAGGCTTCTGGCC
CATGGCGGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGG
5070 APAI, 5097 BALI,

5102 ArgGlyGlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLys
AGAGGAGGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAG
TCTCCTCCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTC
5119 NDEI,

5162 LeuLysLeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAla
CTCAAACCTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCT
GAGTTTGAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGA
5180 NOTI, 5181 EAGI XMA3, 5188 BALI, 5192 PVU2,

5222 GlyTyrSerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrp
GGCTACAGCGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGGCCCGCTGGATCTGG
CCGATGTCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCGGGGGCGACCTAGACC
5246 DRA3,

5282 PheCysLeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP
TTTTGCCTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGATGAAGG
AAAACGGATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTACTTCC
5301 PSTI, 5331 HGE2,

5342 TTGGGGTAAACACTCCGGCCTAAAAAATCTAGAACCCGAGTCGAC
AACCCCATTTGTGAGGCCGATTTTTTTTTTTTTTTAGATCTTGGGCTCAGCTG
5378 XBAI, 5390 SALI,

092445-4300

FIGURE 12

~~12~~
 66
 Std

PAS
C

C.1 C.2

Key

250

48

64

50

3-6

30

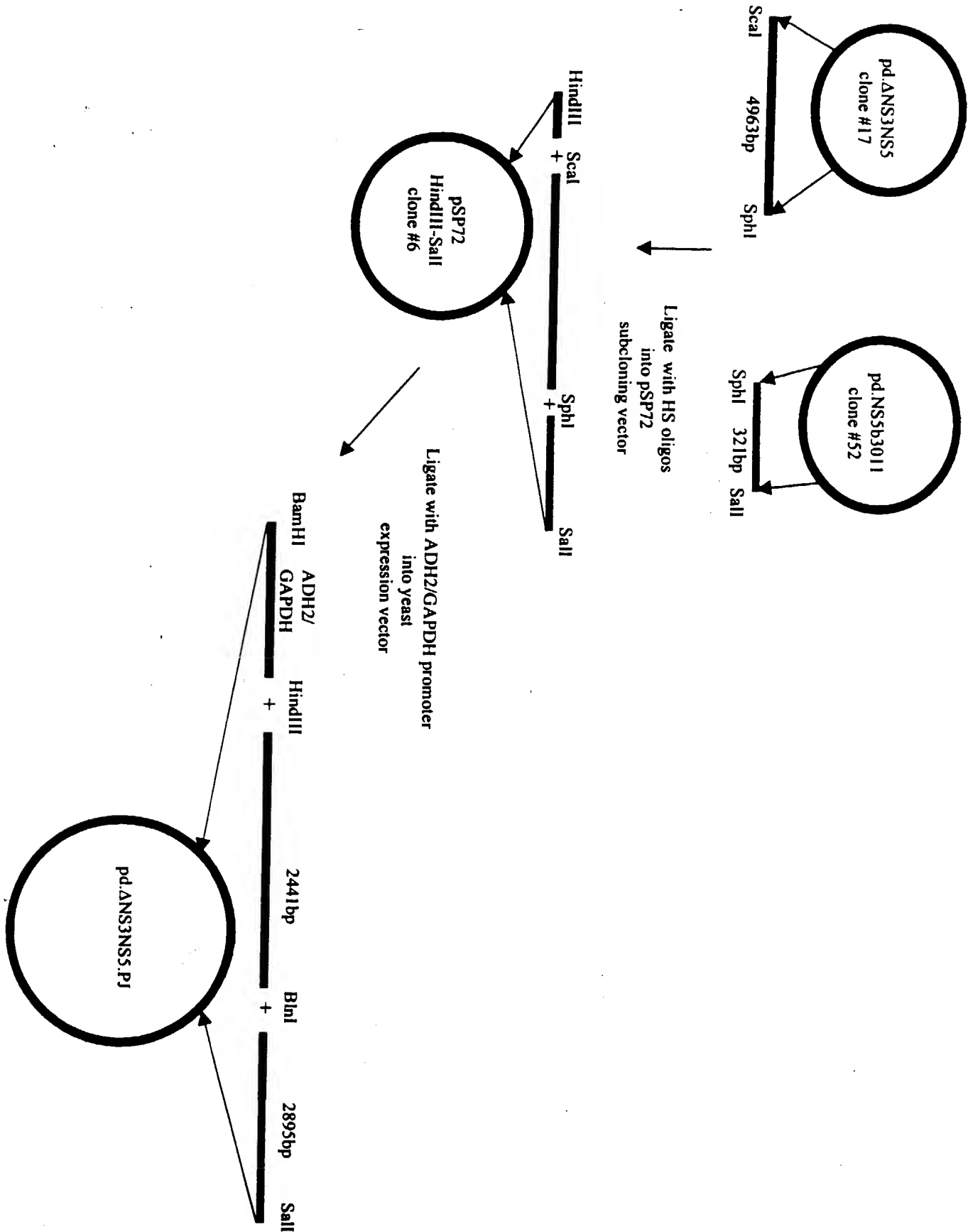
16

6

4

1977-1978 (continued)		1978-1979		1979-1980		1980-1981		1981-1982		1982-1983		1983-1984		1984-1985		1985-1986		1986-1987		1987-1988		1988-1989		1989-1990		1990-1991		1991-1992		1992-1993		1993-1994		1994-1995		1995-1996		1996-1997		1997-1998		1998-1999		1999-2000		2000-2001		2001-2002		2002-2003		2003-2004		2004-2005		2005-2006		2006-2007		2007-2008		2008-2009		2009-2010		2010-2011		2011-2012		2012-2013		2013-2014		2014-2015		2015-2016		2016-2017		2017-2018		2018-2019		2019-2020		2020-2021		2021-2022		2022-2023		2023-2024		2024-2025		2025-2026		2026-2027		2027-2028		2028-2029		2029-2030		2030-2031		2031-2032		2032-2033		2033-2034		2034-2035		2035-2036		2036-2037		2037-2038		2038-2039		2039-2040		2040-2041		2041-2042		2042-2043		2043-2044		2044-2045		2045-2046		2046-2047		2047-2048		2048-2049		2049-2050		2050-2051		2051-2052		2052-2053		2053-2054		2054-2055		2055-2056		2056-2057		2057-2058		2058-2059		2059-2060		2060-2061		2061-2062		2062-2063		2063-2064		2064-2065		2065-2066		2066-2067		2067-2068		2068-2069		2069-2070		2070-2071		2071-2072		2072-2073		2073-2074		2074-2075		2075-2076		2076-2077		2077-2078		2078-2079		2079-2080		2080-2081		2081-2082		2082-2083		2083-2084		2084-2085		2085-2086		2086-2087		2087-2088		2088-2089		2089-2090		2090-2091		2091-2092		2092-2093		2093-2094		2094-2095		2095-2096		2096-2097		2097-2098		2098-2099		2099-2100		2100-2101		2101-2102		2102-2103		2103-2104		2104-2105		2105-2106		2106-2107		2107-2108		2108-2109		2109-2110		2110-2111		2111-2112		2112-2113		2113-2114		2114-2115		2115-2116		2116-2117		2117-2118		2118-2119		2119-2120		2120-2121		2121-2122		2122-2123		2123-2124		2124-2125		2125-2126		2126-2127		2127-2128		2128-2129		2129-2130		2130-2131		2131-2132		2132-2133		2133-2134		2134-2135		2135-2136		2136-2137		2137-2138		2138-2139		2139-2140		2140-2141		2141-2142		2142-2143		2143-2144		2144-2145		2145-2146		2146-2147		2147-2148		2148-2149		2149-2150		2150-2151		2151-2152		2152-2153		2153-2154		2154-2155		2155-2156		2156-2157		2157-2158		2158-2159		2159-2160		2160-2161		2161-2162		2162-2163		2163-2164		2164-2165		2165-2166		2166-2167		2167-2168		2168-2169		2169-2170		2170-2171		2171-2172		2172-2173		2173-2174		2174-2175		2175-2176		2176-2177		2177-2178		2178-2179		2179-2180		2180-2181		2181-2182		2182-2183		2183-2184		2184-2185		2185-2186		2186-2187		2187-2188		2188-2189		2189-2190		2190-2191		2191-2192		2192-2193		2193-2194		2194-2195		2195-2196		2196-2197		2197-2198		2198-2199		2199-2200		2200-2201		2201-2202		2202-2203		2203-2204	
-----------------------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--

FIGURE 13



ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle

662 ACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTTGAGACAATC
TGCACACAGTGGGTCTGTCTCAGCTAAAGTCGGAACCTGGGATGGAAGTGGTAACCTCTGTTAG

ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
722 ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
782 CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTGCACTCGTCC
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG

816 BGLI, 833 DRD1,

ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
842 GTCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGGGAGACT
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGGCTCTGA

881 SACI,

ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGGTAGAA

931 SMAI XMAI,

GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG
CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC

985 STUI,

ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
1022 ACAAAGCAGAGTGGGGAGAACCTTCCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
TGTTTCGTCTCACCCCTCTTGGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA

1069 DRA3,

ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
1082 AGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTGCGCTCAAG
TCCCGAGTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCAAACTAAGCGGAGTTC

ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGGCGCTGTTCAGAATGAAATC
GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG

1150 NCOI,

ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTGCGCCGACCTGGAGGTC
TGGGACTGCGTGGGTCAAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
1262 GTCACGAGACCTGGGTGCTCGTTGGCGGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
 1322 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA
 AGTTGTCCGACGCACAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT
 1369 NAEI,
 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
 1382 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGAGAGTCGTGAAT
 1385 DRD1,
 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAGCAGAAGGCCCTCGGCCTC
 GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG
 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCCTGCTGTCCAGACCAACTGGCAA
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
 1502 PSTI, 1507 TTH3I,
 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
 1562 AAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC
 1565 XHOI, 1586 NDEI,
 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
 1622 GCGGGCTTGTCACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
 CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACTACCGAAAATGTCTGA
 1643 BSTE2, 1677 ALWN1 PVU2,
 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
 1682 GCTGTCAACAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG
 CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC
 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
 1742 GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
 CACCGACGGGTGAGCGGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
 1794 ESP1,
 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
 1802 GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGACGGGTAT
 CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATTA
 1802 KAS1 NARI,
 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
 1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
 1878 SACI, 1899 BSPH1,

ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
 1922 ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCCTAGTCGGC
 TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
 1928 TTH31,
 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
 1982 GTGGTCTGTGCAGCAATACTGCGCCGGCAGCTTGGCCCCGGGCGAGGGGGCAGTGCAGTGG
 CACCAGACACGTCGTTATGACGGGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC
 2004 NAEI, 2017 SMAI XMAI,
 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
 2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCCACGCACTACGTG
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGTGCGTGATGCAC
 2067 SMAI XMAI, 2093 DRA3,
 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
 2102 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
 GGCCTCTCGCTACGTGACGGGGCGCAGTGACGGTATGAGTCGTGGAGTGACATTGGGTC
 2115 PVU2, 2159 ALWN1,
 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
 2162 CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC
 GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
 2164 MST2, 2220 ECON1,
 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
 2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA
 ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT
 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
 2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT
 TTTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCCGCGCCCAT
 2285 ESP1, 2300 PVU2, 2310 BAMHI,
 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
 2342 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
 TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG
 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
 2402 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG
 TGACCTGTACAGTTTTTGGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTGTGAC
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,
 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
 2462 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGGCCCTGTACCCCCCTTCCTGCG
 ACCTCACCTGGAAGGGGTAAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC
 2480 ASE1, 2497 APAI,
 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln

2522 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCGG
GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC
2553 PSTI,
ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
2582 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCGTGCCAG
CACCCCCTGAAGGTGATGCACTGCCCATCTGATGACTGTTAGAATTTACGGGCGACGGTC
2594 DRA3,
ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
2642 GTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCC
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG
ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
2702 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTGAGAGTAGGACTCCACGAATACCCG
GGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC
2757 HGIE2,
ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
2762 GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC
CATCCCAGCGTTAATGGAACGCTCGGGCTTGCCCTGCACCGGCACAACCTGCAGGTACGAG
2809 AAT2,
ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
2822 ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGCGGCGAAGGTTGGCGAGGGGATCACCC
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGCGCGCTTCCAACCGCTCCCTAGTGGG
2850 EAG1 XMA3,
ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
2882 CCGTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
GGGAGACACCGGTGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCGGTTGAACG
2889 BALI, 2903 NHEI,
ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
2942 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGTTGGAGGATACCTCCGTC
2966 ESP1, 2969 SACI,
GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
3002 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG
3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC
3096 BGL2,
ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
3122 CGGAAGTCTCGGAGATTGCCCCAGGCCCTGCCCGTTTGGGCGCGGCCGACTATAACCCC

FIGURE 14 - Page 6

GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG

3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACTGTGGTCCATGGCTGCCCC
GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAAGTACCGACGGGC

3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
CTTCCACCTCCAAAGTCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCTCTC
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG

3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
TGCCCCCGGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA

3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG
CCCCTAGGCCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
GATGTGCTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTACCCACAAT
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTGCTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTTCAAGTGTAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGCGG
TCTGACGTTCAAGACCTGTCGGTAATGGTCTGTCATGAGTTCTCTCAATTTCTGTCGCCG

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG
3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG
3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTGCGTAAG
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
CCAGCTCGTCTCATCGTGTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG
GGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACACGCTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
TACGACGTGGTTACAAAGCTCCCTTGCGCGTGATGGGAAGCTCCTACGGATTCCAATAC
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGAAGTCCAAGAAAACCCCAATG
AGTGGTCTGTGCGCCCACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGGGGTTAC
4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC
4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCTGGCCATCAAGTCC
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTGCGGCGCACCGGTAGTTACAG
4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAGTGGCGC
GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCGC
4345 APAI,

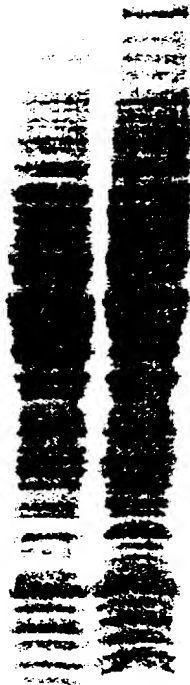
4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
TATCGCAGGTGCCGCGGAGCGGCGTACTGACAAGTAGCTGTGGTAACACCCTCACTTGC
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTTGTGGGAGTGAACG

TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
 4442 TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG
 ATGTAGTTCCGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC
 ^
 4452 SMAI XMAI,
 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
 4502 TGTGGCGACGACTTAGTCTGTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCGCGGAGC
 ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG
 ^ ^
 4508 DRD1, 4511 TTH3I,
 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
 4562 CTGAGAGCCTTCACGGAGGCTATGACCAGGTAATCCGCCCCCCTGGGGACCCCCACAA
 GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT
 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
 4622 CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCCACGAC
 GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG
 ^
 4637 SACI,
 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
 4682 GCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT
 CCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA
 ^
 4731 NRUI,
 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
 4742 GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT
 CGCACCCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA
 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
 4802 GCGGCGACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC
 CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG
 ^ ^
 4806 PFLM1, 4807 DRA3,
 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu
 4862 AGGGACAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA
 TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTT
 ^
 4893 BGL2,
 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
 4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTCTTCACTCCAC
 GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG
 ^
 4954 NCOI,
 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
 4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGGTACCG
 TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC
 ^
 5015 SPHI, 5035 KPNI,
 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly

5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT
5064 APAI, 5091 BALI,
GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT
5113 NDEI,
LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC
GAGTGAGGTTATCGCCGCGACCGGTTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG
5174 NOTI, 5175 EAGI XMA3, 5182 BALI, 5186 PVU2,
SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCCGGCCCCGCTGGATCTGGTTTTGC
TCGCCCCCTCTGTAAATAGTGTGCGCACAGAGTACGGGGCCGGGGCGACCTAGACCAAAACG
5240 DRA3,
LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGATGAATAGTCGAC
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTGGCTACTTATCAGCTG
5295 PSTI, 5336 SALI,

CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA

FIGURE 15



SCANNED, # 14

FIGURE 16 - Page 1

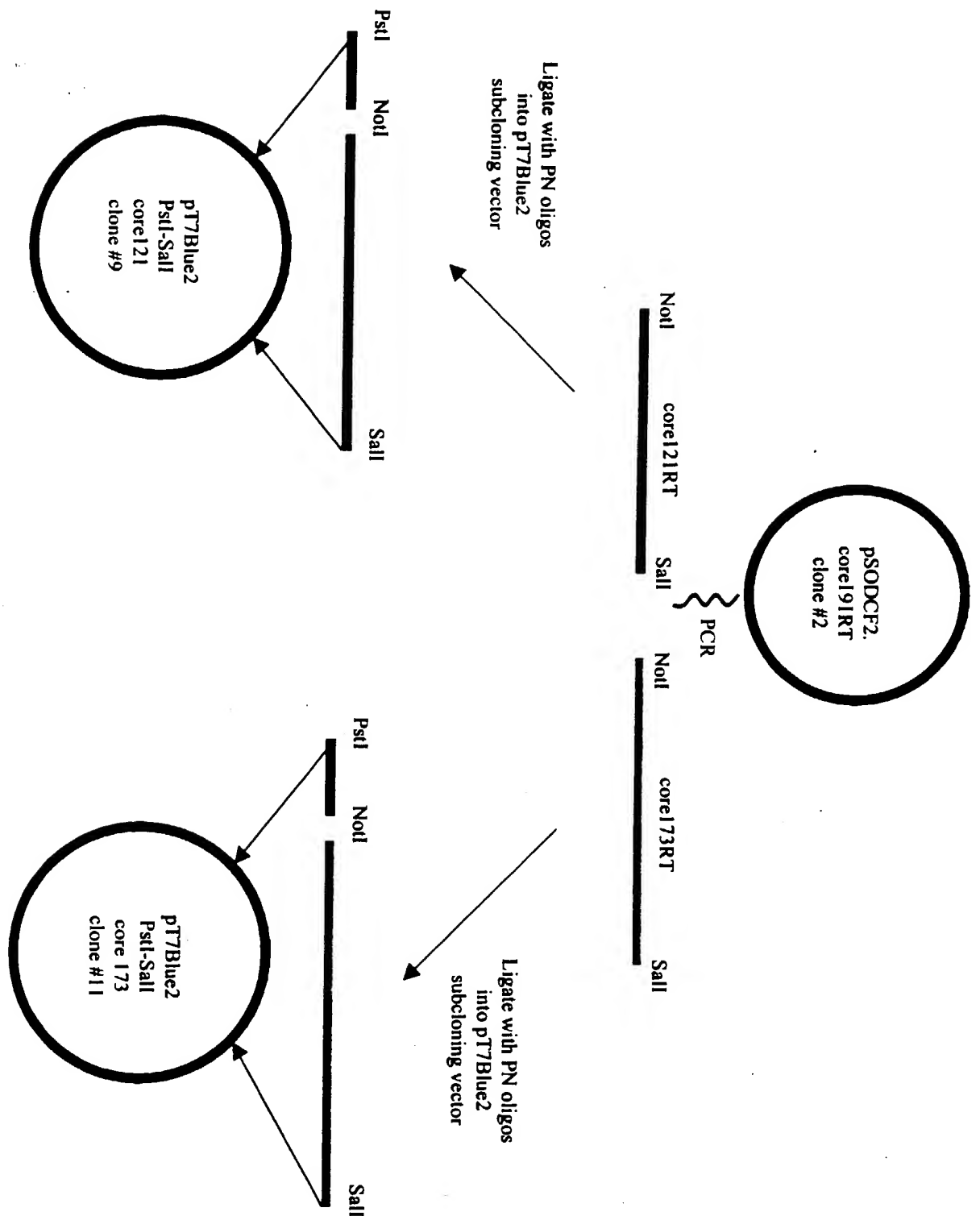
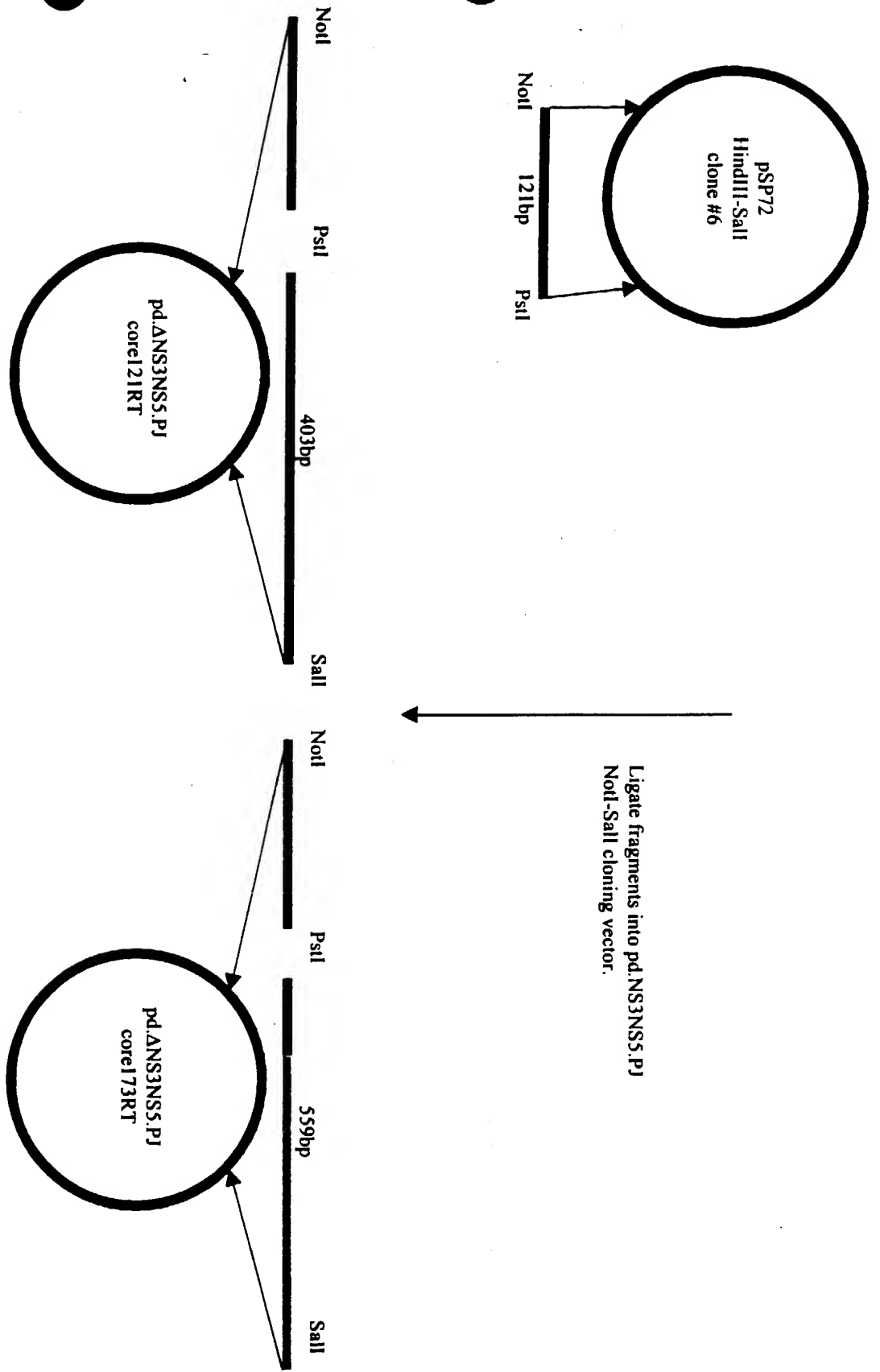


FIGURE 16 - Pa



Ligate fragments into pd.ΔNS3NS5.PJ
NotI-SalI cloning vector.

[illegible]

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn

2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG
^ ^ ^
1 HIND3, 24 NDEI, 52 SCAI,

ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp

62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA
^
116 CLAI,

ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr

122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG

TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys

182 TACGGCAAGTTCTTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAAACA

AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln

242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCGTAACCGTGACAGGAAC TGSTT

AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal

302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG
^
303 ALWN1,

ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe

362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGCCCTCTCTAGGGAAAA

TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis

422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTTCAT
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCCCCTCTGTAGAGTAGAAGACAGTA

FIGURE 17 - Page 2

SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
 482 TCAAAGAAGAAGTGCACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG
 AGTTTCTTCTTACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
 542 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGCGATGTTGTCTGTCGTG
 CGGATGATGGCGCCAGAAGTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC
 ^ ^
 550 SAC2, 560 DRD1,

 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
 602 GCAACCGATGCCCTCATGACCGGTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA
 ^
 615 BSPH1,

 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
 662 ACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC
 TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
 722 ACGTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG
 TGCAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
 782 CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG
 ^ ^
 816 BGLI, 833 DRD1,

 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
 842 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCGAGACT
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA
 ^
 881 SACI,

 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
 902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
 TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA
 ^
 931 SMAI XMAI,

 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
 962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG
 CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC
 ^
 985 STUI,

 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
 1022 ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
 TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA
 ^
 1069 DRA3,

 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
 1082 AGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG

TCCCGAGTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTC

ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
 1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGTAAATC
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG
 ^

1150 NCOI,

ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
 1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG
 ^ ^ ^ ^ ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
 1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG
 CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
 1322 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA
 AGTTGTCCGACGCACCAAGTATCACCCGTCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT
 ^

1369 NAEI,

ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
 1382 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGAGTCGTGAAT
 ^

1385 DRD1,

ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC
 GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCTGCTGTCCAGACCAACTGGCAA
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
 ^ ^

1502 PSTI, 1507 TTH3I,

LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
 1562 AAATCGAGACCTTCTGGGCGAAGCATATGTGGAAGTTCATCAGTGGGATACAATACTTG
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCATGTTATGAAC
 ^ ^

1565 XHOI, 1586 NDEI,

AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
 1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
 CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACTACCGAAAATGTCTGA
 ^ ^

1643 BSTE2, 1677 ALWN1 PVU2,

AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
 1682 GCTGTACACGACCCACTAACCCTAGCCAAACCCCTCCTCTTCAACATATTGGGGGGGTGG
 CGACAGTGGTTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
1742 GTGGCTGCCCAGCTCGCCGCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAACACCCGCGACCGAATCGA
1794 ESP1,
GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
1802 GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT
1802 KAS1 NARI,
GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
1878 SACI, 1899 BSPH1,
ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
1922 ACGGAGGACCTGGTCAATCTACTGCCCCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
1928 TTH3I,
ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCCGGGCGAGGGGGCAGTGCACTGG
CACCAGACACGTCGTTATGACGCGGGCGGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC
2004 NAEI, 2017 SMAI XMAI,
MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGTGCGTGATGCAC
2067 SMAI XMAI, 2093 DRA3,
ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
2102 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC
2115 PVU2, 2159 ALWN1,
LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
2162 CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC
GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
2164 MST2, 2220 ECON1,
TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTGAGCGACTTTAAGACCTGGCTA
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT
LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT
TTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTGCAGCCCCATA
2285 ESP1, 2300 PVU2, 2310 BAMHI,

003474660

2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
 AAGGGGGTCTGGCGAGGGGATGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
 TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

 2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG
 TGACCTGTACAGTTTTTGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGTA
 ^ ^ ^
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

 2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCCTGTACCCCCCTTCCTGCG
 ACCTCACCCCTGGAAGGGGTAATTACGGATGTGGTGGCCGGGGACATGGGGGGAAGGACGC
 ^ ^
 2480 ASE1, 2497 APAI,

 2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG
 GGCTTGATGTGCAAGCGGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC
 ^
 2553 PSTI,

 2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCGTGCCAG
 CACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC
 ^
 2594 DRA3,

 2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
 GTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCC
 CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGGGATGTATCCAAACGCGGG

 2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTGAGAGTAGGACTCCACGAATACCCG
 GGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC
 ^
 2757 HGIE2,

 2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
 GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC
 CATCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACCTGCAGGTACGAG
 ^
 2809 AAT2,

 2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
 ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGCGAAGGTTGGCGAGGGGATCACCC
 TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCTAGTGGG
 ^
 2850 EAG1 XMA3,

 2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
 GGGAGACACCGGTGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG
 ^ ^
 2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGTTGGAGGATACCTCCGTC

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGTTGATTCTGGACTCC
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
 AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC

3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
 CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC
 GCCTTCAGAGCCTCTAAGCGGGTCCCGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG

3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG
 GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGG

3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
 CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCCTC
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
 ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG

3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC
 AGTTGAAGGCCGTAATGCCCCGTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
 TGCCCCCGGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT
 ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA

3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG
 CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGTTTGGCCTC

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
 GATGTGCTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGGCGC
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTTCAGTGTAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTTCGTCGCCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCCGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGTCGTAAG
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTG

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
CCAGCTCGTCTCATCGTGTTCCTCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
TACGACGTGGTTACAAAGCTCCCCTTGCCGTGATGGGAAGCTCCTACGGATTCCAATAC
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG
AGTGGTCCTGTGCGCCCACTTAAGGAGCACGTTCCGCACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCACTGACTCTCGCTGTAGGCATGC

FIGURE 17 - Page 8

4229 DRD1, 4236 ALWN1,

GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
4262 GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGGTGGCCATCAAGTCC
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTTCGGGCGCACCGGTAGTTCAGG

4301 BGLI, 4308 BALI,

LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
4322 CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC
GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG

4345 APAI,

TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
4382 TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACCTAGCTGTGGTAACACCCTCACTTGC
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
4442 TACATCAAGGCCCCGGGCAGCCTGTGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG
ATGTAGTTCGGGGCCCGTTCGGACAGCTCGGCCTCCCGAGGTCTGACGTGGTACGAGCAC

4452 SMAI XMAI,

CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
4502 TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCTGCGCCGCTCG

4508 DRD1, 4511 TTH3I,

LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
4562 CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT

ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
4622 CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTGAGTCGCCCCACGAC
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG

4637 SACI,

GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
4682 GGCCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA

4731 NRUI,

AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
4742 GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT
CGCACCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA

AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
4802 GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC
CGGGGGTGTGACACCCGCTCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

Downloaded from www.jstor.org

FIGURE 17 - Page 9

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA
TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT
^

4893 BGL2,

ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAAGTGAGGTG
^

4954 NCOI,

SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAAGTTGGGGTACCG
TCAATGAGAGGTCCACTTTAGTTATCCCACCGGCGTACGGAGTCTTTGAACCCCATGGC
^

5015 SPHI, 5035 KPNI,

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
5042 CCCTTGCGAGCTTGGAGACACCGGGCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
GGGAACGCTCGAACCTCTGTGGCCCGGCCCTCGCAGGCGCGATCCGAAGACCGGTCTCTCT
^

5064 APAI, 5091 BALI,

GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA
CCGTCCCGACGGTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT
^

5113 NDEI,

LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
5162 CTCCTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTACGGCTGGCTAC
GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG
^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCCGGCCCCGCTGGATCTGGTTTTGC
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAACG
^

5240 DRA3,

LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn
5282 CTACTCCTGCTTGCTGCGAGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTGGCTTACTCGTGCTTA
^

5295 PSTI,

ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe
5342 CCTAAACCTCAAAGAAAGACCAACGTAACACCAACCGGCGCCGAGGACGTCAAGTTC
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCCGCGCGTCTGCAGTTCAAG
^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
5402 CCGGGTGGCGGTGAGATCGTTGGTGGAGTTTACTTGTGTCGCGCAGGGGCCCTAGATTG
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC
^

0974764260

FIGURE 17 - Page 10

5449 APAI,

GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
 5462 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCGAACCTCGAGGTAGACGTCAGCCT
 CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
 5522 ATCCCCAAGGCTCGTCGGCCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC
 TAGGGGTTCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
 5582 CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG
 GAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysOC AM
 5642 CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTGCGCAATTTGGGTAAGTAATAGTCG
 GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTATTATCAGC

5650 APAI, 5698 SALI,

5702 AC
 TG

```

# 1. Import the data
data = pd.read_csv('data.csv')

# 2. Data cleaning
data = data.dropna()
data = data[data['age'] > 0]

# 3. Feature engineering
data['age_group'] = pd.cut(data['age'], bins=[0, 18, 35, 50, 65, 80],
                           labels=['Young', 'Young Adult', 'Adult', 'Middle-aged', 'Old'])

# 4. Model training
X = data[['age', 'income', 'education']]
y = data['target']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
                                                    random_state=42)

# 5. Model evaluation
model = LogisticRegression()
model.fit(X_train, y_train)

y_pred = model.predict(X_test)

# 6. Performance metrics
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)
f1_score = f1_score(y_test, y_pred)

print(f'Accuracy: {accuracy}, Precision: {precision}, Recall: {recall}, F1 Score: {f1_score}')

```

2

FIGURE 18 - Page 2

ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
 662 ACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC
 TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
 722 ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG
 TCGGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCCTCCCCCTTC

 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
 782 CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTCTGACTCGTCC
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG
 ^ ^
 816 BGLI, 833 DRD1,

 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
 842 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCCGAGACT
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGGCTCTGA
 ^
 881 SACI,

 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
 902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
 TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTCGGTAGAA
 ^
 931 SMAI XMAI,

 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
 962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTCTATCCAG
 CTTAAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC
 ^
 985 STUI,

 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
 1022 ACAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
 TGTTTCGTCTCACCCCTCTTGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA
 ^
 1069 DRA3,

 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
 1082 AGGGCTCAAGCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTCGCCTCAAG
 TCCCGAGTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTC

 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
 1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGTAAATC
 GGGTGGGAGGTACCCGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG
 ^
 1150 NCOI,

 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
 1202 ACCCTGACGCACCCAGTCACCAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG
 ^ ^ ^ ^ ^
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
 1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCCTG

FIGURE 18 - Page 3

CAGTGTCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGGCGCATAACGGAC
 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
 1322 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA
 AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT
 1369 NAEI,
 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
 1382 CCTGACAGGGAAGTCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGAGAGTCGTGAAT
 1385 DRD1,
 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC
 GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG
 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
 1502 PSTI, 1507 TTH3I,
 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
 1562 AAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC
 1565 XHOI, 1586 NDEI,
 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
 1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
 CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACTACCGAAAATGTGCA
 1643 BSTE2, 1677 ALWN1 PVU2,
 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
 1682 GCTGTCAACAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG
 CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC
 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
 1742 GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
 CACCGACGGGTGAGCGGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
 1794 ESP1,
 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
 1802 GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGAGGGGTAT
 CCGCGGCGGTAGCCGTACAACTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATTA
 1802 KAS1 NARI,
 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
 1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
 1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
ACGGAGGACCTGGTCAATCTACTGCCCCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG
CACCAGACACGTCGTTATGACGCGGGCGTGCAACCGGGGCCGCTCCCCCGTCACGTCACC
2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCCTGATGCAC
2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
CCGGAGAGCGATGCAGCTGCCCCGCTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC
2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTAACCACTCCATGCTCCGGTTCC
GAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT
TTTCGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCATTA
2285 ESP1, 2300 PVU2, 2310 BAMHI,

2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG
TGACCTGTACAGTTTTTGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATAC
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGGCCCCTGTACCCCCCTTCTGCG
ACCTCACCCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGACATGGGGGGAAGGACGC
2480 ASE1, 2497 APAI,

CCCTTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC

FIGURE 18 - Page 5

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG
 GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC
 2553 PSTI,
 2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG
 CACCCCTGAAGGTGATGCACTGCCCATCTGATGACTGTTAGAATTTACGGGCACGGTC
 2594 DRA3,
 2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
 GTCCCATCGCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC
 CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG
 2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG
 GGGACGTTTCGGGAACGACGCCCTCTCCATAGTAAGTCTCATCTGAGGTGCTTATGGGC
 2757 HGIE2,
 2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
 GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC
 CATCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACGTGAGGTACGAG
 2809 AAT2,
 2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
 ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGCGAAGGTTGGCGAGGGGATCACCC
 TGACTIONGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCTAGTGGG
 2850 EAG1 XMA3,
 2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
 GGGAGACACCGGTTCGAGGAGCCGATCGGTTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG
 2889 BALI, 2903 NHEI,
 2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
 ACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGTTGGAGGATACCTCCGTC
 2966 ESP1, 2969 SACI,
 3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG
 3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
 AAGCTAGGCGAACACCGCCTCTCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC
 3096 BGL2,
 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro

3122 CGGAAGTCTCGGAGATTCCGGACAGGCCCTGCCCGTTTGGGGCGCGCGGACTATAACCCC
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGCCTGATATTGGGG
3143 ALWN1, 3164 EAG1 XMA3,
ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
3182 CCGCTAGTGGAGACGTGGAAAAAGCCCCGACTACGAACCACTGTGGTCCATGGCTGCCCCG
GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC
3217 HGIE2, 3229 NCOI,
LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
3242 CTTCCACCTCCAAAGTCCCCCTCTGTGCCTCCGCCCTCGGAAGAAGCGGACGGTGGTCCCTC
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG
ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
3302 ACTGAATCAACCTTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG
3332 SACI, 3346 HIND3,
SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
3362 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCCCCCCTTCTGGC
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG
CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
3422 TGCCCCCCCCGACTCCGACGCTGAGTCTTATCCTCCATGCCCCCCTGGAGGGGGAGCCT
ACGGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
3437 EAM11051,
GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
3482 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG
CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,
AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
3542 GATGTGCGTGTGCTGCTCAATGTCTTACTCTTGACAGGCGCACTCGTCACCCCGTGCGCC
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG
3589 DRA3, 3600 SAC2,
AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
3602 GCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTACACACAAT
CGCCTTCTTGCTCTTTGACGGGTAGTTACGTGATTGCTTGAGCAACGATGCAGTGGTGTTA
3611 ALWN1, 3655 PFLM1,
LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAAGATCATTGAC
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAACCTG
3681 DRA3,
ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
3722 AGACTGCAAGTTCTTGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
TCAAAAGTGAAGGCTAAGTTCCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
3902 GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
4022 CCAGCTCGTCTCATCGTGTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG
GGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG
AGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGGGGTTAC

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCTGGCCATCAAGTCC
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTCAGG

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC
GAGTGGCTCTCCGAAATACAACCCCGGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
TATCGCAGGTGCCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
 4442 TACATCAAGGCCCGGGCAGCTTGTCTGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG
 ATGTAGTTCCGGGCGCGTCTGGACAGCTCGGCGTCCCGAGGTCTGACGTGGTACGAGCAC
 4452 SMAI XMAI,
 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
 4502 TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCGGCGAGC
 ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTCTGCGCCGCTCG
 4508 DRD1, 4511 TTH3I,
 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
 4562 CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA
 GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT
 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
 4622 CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTGAGTCGCCCCACGAC
 GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG
 4637 SACI,
 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
 4682 GGCGCTGGAAAGAGGGTCTACTACCTACCCGTGACCCTACAACCCCCCTCGCGAGAGCT
 CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA
 4731 NRUI,
 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
 4742 GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCTTGCTAGGCAACATAATCATGTTT
 CGCACCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA
 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
 4802 GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC
 CGGGGGTGTTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG
 4806 PFLM1, 4807 DRA3,
 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu
 4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGGAGATCTACGGGGCCTGCTACTCCATAGAA
 TCCCTGGTTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT
 4893 BGL2,
 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
 4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
 GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG
 4954 NCOI,
 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
 4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAACTTGGGGTACCG
 TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC
 5015 SPHI, 5035 KPNI,

FIGURE 18 - Page 9

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
 CCCTTGGCAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
 GGAACGCTCGAACCTCTGTGGCCCGGGCTCGCAGGCGCGATCCGAAGACCGGTCTCCT
 5064 APAI, 5091 BALI,
 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
 5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA
 CCGTCCCGACGGTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT
 5113 NDEI,
 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
 5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC
 GAGTGAGGTTATCGCCGGCGACCGGTGACCTGAACAGGCCGACCAAGTGCCGACCGATG
 5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,
 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
 5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGGCCCGCTGGATCTGGTTTTGC
 TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGGCCGGGGCGACCTAGACCAAACG
 5240 DRA3,
 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn
 5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT
 GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGTTGGCTTACTCGTGCTTA
 5295 PSTI,
 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe
 5342 CCTAAACCTCAAAGAAAGACCAACGTAACACCAACCGGCGGCCGAGGACGTCAAGTTC
 GGATTTGGAGTTTCTTTCTGGTTTGCAATTGTGGTTGGCCCGCGCGCTCCTGCAGTTCAAG
 5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,
 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
 5402 CCGGGTGGCGGTGAGATCGTTGGTGGAGTTTACTTGTGTCGCGCAGGGGCCCTAGATTG
 GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC
 5449 APAI,
 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
 5462 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTACAGCT
 CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCCGA
 5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,
 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
 5522 ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC
 TAGGGGTTCCGAGCAGCCGGGCTCCCGTCTGGACCCGAGTCGGGCCCATGGGAACCGGG
 5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,
 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
 5582 CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG
 GAGATACCGTTACTCCCGACGCCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

FIGURE 18 - Page 10

ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp
 5642 CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTGCGCGCAATTTGGGTAAGGTCATCGAT
 GGATCGACCCCGGGGTGTCTGGGGGCCGATCCAGCGCGTTAAACCCATTCCAGTAGCTA

5650 APAI, 5696 CLAI,

ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu
 5702 ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCTCTT
 TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA

5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,

GlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAspGlyValAsnTyr
 5762 GGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTCTGGAAGACGGCGTGAACTAT
 CCTCCGCGACGGTCCCGGACCGGTACCGCAGGCCCAAGACCTTCTGCCGCACTTGATA

5772 BSTXI, 5775 APAI,

AlaThrGlyAsnLeuProGlyCysSerOC AM
 5822 GCAACAGGGAACCTTCCTGGTTGCTCTTAATAGTCGAC
 CGTTGTCCCTTGGAAGGACCAACGAGAATTATCAGCTG

5854 SALI,

5854 SALI,

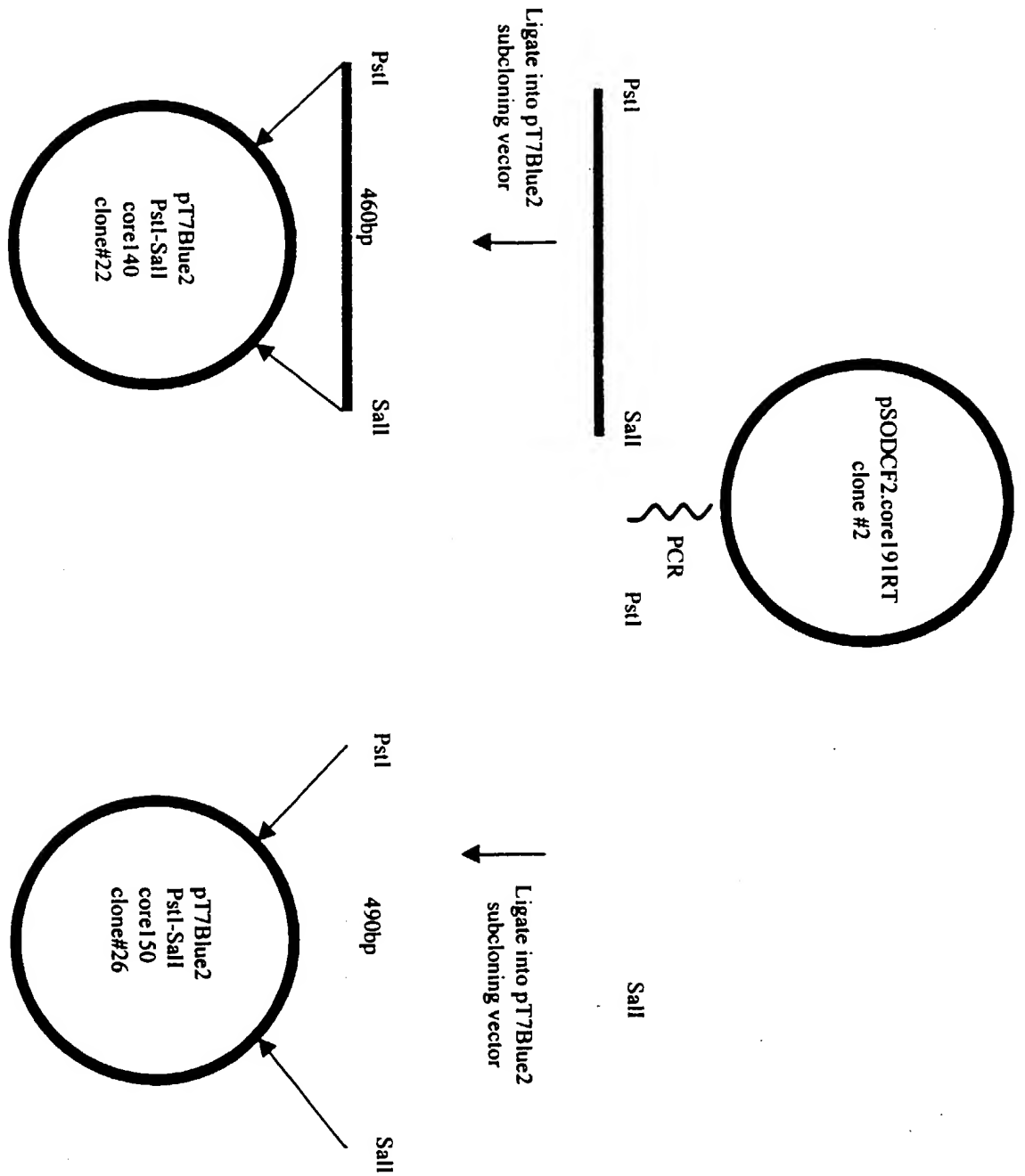


FIGURE 20 - Pa

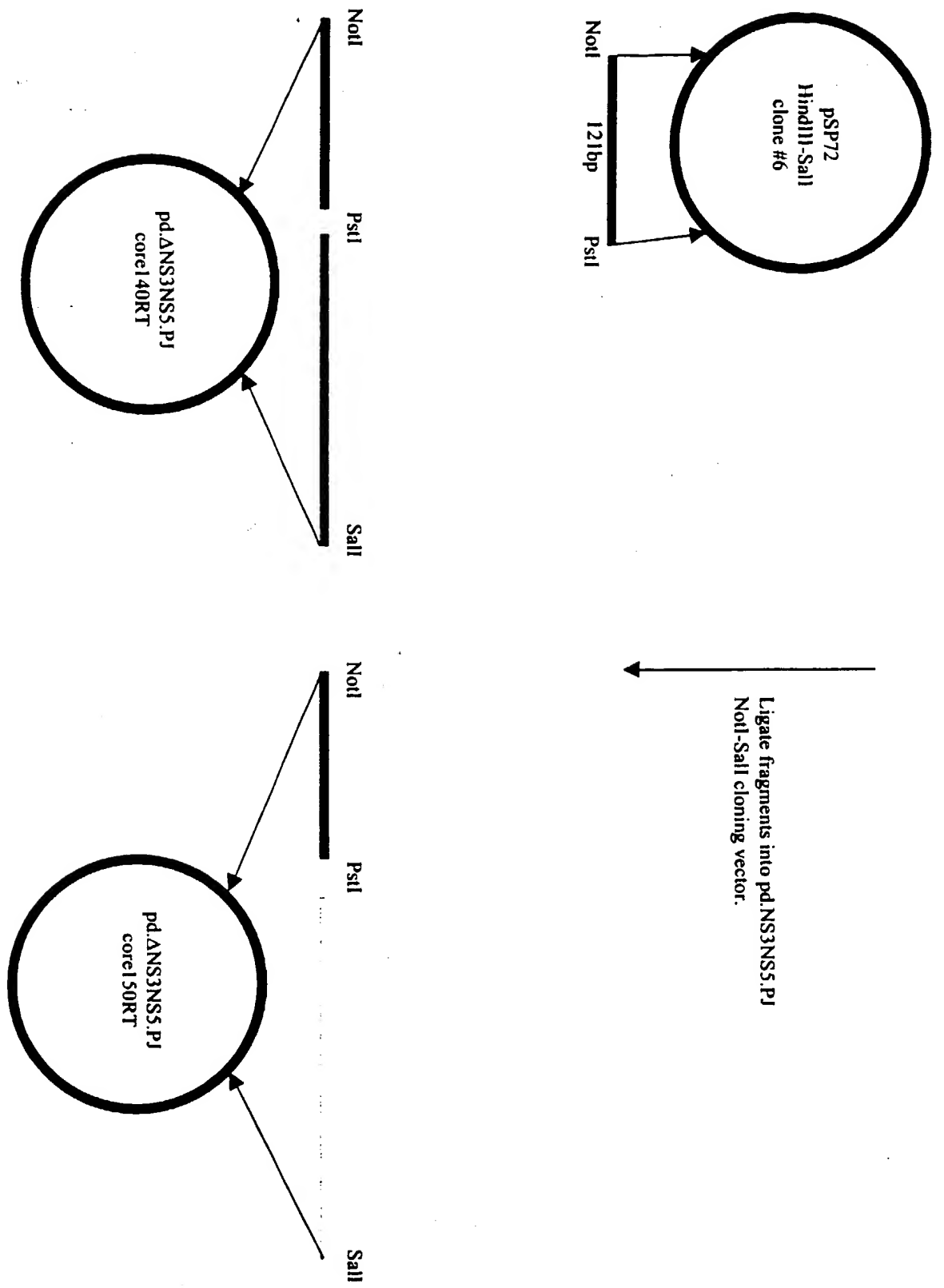


FIGURE 21 - Page 1

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
 2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
 TCGAATGTTTTGTTTTACCGACGTATACGTTCGAGTCCCGATATTCCACGATCATGAGTTG
 ^ ^ ^
 1 HIND3, 24 NDEI, 52 SCAI,
 ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
 62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
 GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCCTAGCTA
 ^
 116 CLAI,
 ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
 122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCAGTACTCCACC
 GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG
 TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
 182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT
 ATGCCGTTCAAGGAACGGCTGCCGCCCCACGAGCCCCCGGAATACTGTATTATTAAACA
 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
 242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA
 CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACTGGTT
 AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
 302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC
 CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG
 ^
 303 ALWN1,
 ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
 362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGAGAGATCCCTTTT
 TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAA
 TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
 422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTTCAT
 ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
 482 TCAAAGAAGAAGTGGCAGCAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG
 AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
 542 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTGCGTG
 CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC
 ^ ^
 550 SAC2, 560 DRD1,

 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
 602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA
 ^
 615 BSPH1,

 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
 662 ACGTGTGTCAACCCAGACAGTCGATTTTACGCTTGACCCTACCTTCACCATTGAGACAATC
 TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAAGTCTGTTAG

 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
 722 ACGTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG
 TCGCAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
 782 CCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCCCTCCGGCATGTTTCGACTCGTCC
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG
 ^ ^
 816 BGLI, 833 DRD1,

 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
 842 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCAGCCCCGCCGAGACT
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA
 ^
 881 SACI,

 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
 902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
 TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA
 ^
 931 SMAI XMAI,

 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
 962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCCTTTCTATCCCAG
 CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC
 ^
 985 STUI,

 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
 1022 ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
 TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA
 ^
 1069 DRA3,

 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
 1082 AGGGCTCAAGCCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG

TCCCGAGTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTC

ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
 1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGATGAAATC
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG
 ^
 1150 NCOI,

ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
 1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG
 ^ ^ ^ ^ ^
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
 1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG
 CAGTGTCTGCTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
 1322 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA
 AGTTGTCCGACGCACAGTATCACCCGTCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT
 ^
 1369 NAEI,

ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
 1382 CCTGACAGGGAAGTCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGAGTCGTGAAT
 ^
 1385 DRD1,

ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAGCAGAAAGGCCCTCGGCCTC
 GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
 ^ ^
 1502 PSTI, 1507 TTH3I,

LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
 1562 AAATCGAGACCTTCTGGGCGAAGCATATGTGGAACCTCATCAGTGGGATACAATACTTG
 TTTGAGCTCTGGAAGACCCGTTCTGATACACCTTGAAGTAGTCACCCTATGTTATGAAC
 ^ ^
 1565 XHOI, 1586 NDEI,

AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
 1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
 CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACACCGAAAATGTCTGA
 ^ ^
 1643 BSTE2, 1677 ALWN1 PVU2,

AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
 1682 GCTGTACACGACCCACTAACCCTAGCCAAACCCCTCCTCTTCAACATATTGGGGGGGTGG
 CGACAGTGGTCCGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

09447-647260

ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
 1742 GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
 CACCGACGGGTGAGCGGGCGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
 ^
 1794 ESP1,
 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
 1802 GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGAGGGGTAT
 CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT
 ^
 1802 KAS1 NARI,
 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
 1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC
 CCGCGCCCGACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
 ^ ^
 1878 SACI, 1899 BSPH1,
 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
 1922 ACGGAGGACCTGGTCAATCTACTGCCCCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC
 TGCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCC
 ^
 1928 TTH3I,
 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
 1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCACTGG
 CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCGCTCCCCCGTCACGTCACC
 ^ ^
 2004 NAEI, 2017 SMAI XMAI,
 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
 2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC
 ^ ^
 2067 SMAI XMAI, 2093 DRA3,
 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
 2102 CCGGAGAGCGATGCAGCTGCCCGCTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
 GGCCTCTCGCTACGTGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC
 ^ ^
 2115 PVU2, 2159 ALWN1,
 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
 2162 CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTAACCTCCATGCTCCGGTTC
 GAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
 ^ ^
 2164 MST2, 2220 ECON1,
 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
 2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA
 ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT
 ^ ^
 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
 2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT
 TTTTCGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCAT
 ^ ^ ^
 2285 ESP1, 2300 PVU2, 2310 BAMHI,

LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
 2342 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
 TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
 2402 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG
 TGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATC
 ^ ^ ^
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
 2462 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCG
 ACCTCACCTGGAAGGGGTAAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC
 ^ ^
 2480 ASE1, 2497 APAI,

 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
 2522 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG
 GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC
 ^
 2553 PSTI,

 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
 2582 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCGTGCCAG
 CCCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC
 ^
 2594 DRA3,

 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
 2642 GTCCCATCGCCCGAATTTTTACAGAATTGGACGGGGTGCGCTACATAGGTTTGCGCCC
 CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG

 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
 2702 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTGAGAGTAGGACTCCACGAATACCCG
 GGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC
 ^
 2757 HGIE2,

 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
 2762 GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC
 CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAATGCAGGTACGAG
 ^
 2809 AAT2,

 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
 2822 ACTGATCCCTCCCATATAACAGCAGAGGCGCGCGGCGAAGGTTGGCGAGGGGATCACCC
 TGACTAGGGAGGGTATATTGTCGTCTCCGCGCGCCGCTTCCAACCGCTCCCTAGTGGG
 ^
 2850 EAG1 XMA3,

 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
 2882 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
 GGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG
 ^
 2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC
^ ^

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
GAGATGGGCGGCAACATCACCAGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
AAGCTAGGCGAACACCGCCTCCTCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC
^

3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCCGTTGGGCGCGGCCGACTATAACCCC
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGCGCCTGATATTGGGG
^ ^

3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
CCGCTAGTGGAGACGTGGAAAAAGCCCCGACTACGAACCACTGTGGTCCATGGCTGCCCCG
GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC
^ ^

3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
ACTGAATCAACCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG
^ ^

3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
^

3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
GGGGATCCGGATCTTAGCGACGGGTGATGGTCAACGGTCAGTAGTGAGGCCAACCGGGAG
CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC
^ ^ ^

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
GATGTGCTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

00221647460

3589 DRA3, 3600 SAC2,

AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
3602 GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTGCGTGTTA

3611 ALWN1. 3655 PFLM1.

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC
 AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGCTTCTCTTTCAGTGTAAACTG

3681 DRA3,

ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCTCCAATTTTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC
AGTTTTCACTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
3902 GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTTCAGCCTGAGAAGGGGGGTCGTAAG
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
4022 CCAGCTCGTCTCATCGTGTCCCCGATCTGGGCGTGCGCGTGTGCGAAAGATGGCTTTG
GGTCGAGCAGAGTAGACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
TACGACGTGGTTACAAAGCTCCCCCTTGCCGTGATGGGAAGCTCCTACGGATTCCAATAC
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAACCCCAATG
AGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTCGCACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
4202 GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA
TCCCTGGTTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT

4893 BGL2,

ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCTACTCCAC
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAACTTGGGGTACCG
TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC

5015 SPHI, 5035 KPNI,

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
GGGAACGCTCGAACCTCTGTGGCCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCGAGTAAGAACAAAGCTCAAA
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT

5113 NDEI,

5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTACGGCTGGCTAC
GAGTGAGGTTATCGCCGGCAGCCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGGCGGGGCGACCTAGACCAAACG

5240 DRA3,

5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA

5295 PSTI,

ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe
5342 CCTAAACCTCAAAGAAAGACCAACGTAACACCAACCGGCGGCCGAGGACGTCAAGTTC
GGATTGGAGTTTCTTCTGGTTTGCATTGTGGTTGGCCGCCGCGCTCCTGCAGTTCAAG

5380 NOT1, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
5402 CCGGGTGGCGGT CAGATCGTTGGTGGAGTTTACTTGTTGCCGCGCAGGGGCCCTAGATTG
GGCCACCGCCAGTCTAGCAACCACTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCGAACCTCGAGGTAGACGTCAGCCT
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGGCCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCGTGGCTCTCGG
GAGATACCGTTACTCCCGACGCCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp
CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTGCGCAATTTGGGTAAGGTCATCGAT
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValOC AM
ACCCTTACGTGCGGCTTCGCGGACCTCATGGGGTACATACCGCTCGTCTAATAGTCGAC
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGATTATCAGCTG

5724 HGIE2, 5755 SALI,

5449 APAI, 5462 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCGAACCTCGAGGTAGACGTCAGCCT
5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2, 5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI, 5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
5642 CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTGCGCAATTTGGGTAAGGTCATCGAT
5650 APAI, 5696 CLAI, 5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValOC AM
5724 HGIE2, 5755 SALI, 5724 HGIE2, 5755 SALI,

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
 2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
 TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG
 ^
 1 HIND3, 24 NDEI, 52 SCAI,
 ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
 62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
 GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA
 ^
 116 CLAI,
 ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
 122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCAGTACTCCACC
 GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG
 TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
 182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT
 ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATTAAACA
 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
 242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA
 CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGT
 AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
 302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC
 CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG
 ^
 303 ALWN1,
 ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
 362 ACTGTGCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGAGAGATCCCTTTT
 TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA
 TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
 422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTGTCAT
 ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
 482 TCAAAGAAGAAGTGCACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG
 AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC
 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
 542 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTCGTG
 CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC
 550 SAC2, 560 DRD1,
 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
 602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA
 615 BSPH1,
 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
 662 ACGTGTGTACCCAGACAGTCGATTTACGCCTTGACCCTACCTTCACCATTGAGACAATC
 TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACCTCTGTTAG
 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
 722 ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG
 TGCAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC
 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
 782 CCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCCTCCGGCATGTTGACTCGTCC
 GGTCCTGATGCTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG
 816 BGLI, 833 DRD1,
 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
 842 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCCGAGACT
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA
 881 SACI,
 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
 902 ACAGTTAGGCTACGAGCGTACATGAACCCCCGGGCTTCCCGTGTGCCAGGACCATCTT
 TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTCGGTAGAA
 931 SMAI XMAI,
 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
 962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG
 CTTAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC
 985 STUI,
 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
 1022 ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
 TGTTTCGTCTACCCCTCTTGGAAGGAATGGACCATCGCATGTTTCGGTGCCACACGCGA
 1069 DRA3,
 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
 1082 AGGGCTCAAGCCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG

AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
1682 GCTGTCACCAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

FIGURE 22 - Page 4

ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
 1742 GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
 CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
 ^

1794 ESP1,

GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
 1802 GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCTCATAGACATCCTTGCAGGGTAT
 CCGCGGCGGTAGCCGTACAACTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA
 ^

1802 KAS1 NARI,

GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
 1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCTCC
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
 ^

1878 SACI, 1899 BSPH1,

ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
 1922 ACGGAGGACCTGGTCAATCTACTGCCCCGCCATCCTCTCGCCCGAGCCCTCGTAGTCGGC
 TGCCTCTCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
 ^

1928 TTH3I,

ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
 1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGCGAGTGCAGTGG
 CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC
 ^

2004 NAEI, 2017 SMAI XMAI,

MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
 2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC
 ^

2067 SMAI XMAI, 2093 DRA3,

ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
 2102 CCGGAGAGCGATGCAGCTGCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
 GGCTCTCGCTACGTGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC
 ^

2115 PVU2, 2159 ALWN1,

LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
 2162 CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC
 GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
 ^

2164 MST2, 2220 ECON1,

TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
 2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA
 ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
 2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCCTGCCAGCGCGGGTAT
 TTTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCAT
 ^

2285 ESP1, 2300 PVU2, 2310 BAMHI,

```

# Getting the sample size
n = 1000
n1 = 500
n2 = 500

# Getting the data
data = read.csv("data.csv")

# Getting the data for the first group
data1 = data[1:n1,]

# Getting the data for the second group
data2 = data[(n1+1):(n1+n2),]

# Getting the mean for the first group
mean1 = mean(data1[, "value"])

# Getting the mean for the second group
mean2 = mean(data2[, "value"])

# Getting the standard deviation for the first group
sd1 = sd(data1[, "value"])

# Getting the standard deviation for the second group
sd2 = sd(data2[, "value"])

# Getting the t-statistic
t = (mean1 - mean2) / (sqrt((sd1^2 / n1) + (sd2^2 / n2)))

# Getting the p-value
p = pt(t, df = n1 + n2 - 2)

# Getting the confidence interval
ci = c(mean1 - mean2 - qt(0.975, df = n1 + n2 - 2) * sqrt((sd1^2 / n1) + (sd2^2 / n2)),
       mean1 - mean2 + qt(0.975, df = n1 + n2 - 2) * sqrt((sd1^2 / n1) + (sd2^2 / n2)))

# Printing the results
print(p)
print(ci)

```

ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
 2942 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC
 ^ ^
 2966 ESP1, 2969 SACI,
 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
 3002 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAAACAAAGTGGTGATTCTGGACTCC
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG
 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
 3062 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
 AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC
 ^
 3096 BGL2,
 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
 3122 CGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC
 GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG
 ^ ^
 3143 ALWN1, 3164 EAG1 XMA3,
 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
 3182 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG
 GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGG
 ^ ^
 3217 HGIE2, 3229 NCOI,
 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
 3242 CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG
 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
 3302 ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG
 ^ ^
 3332 SACI, 3346 HIND3,
 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
 3362 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCGCCCTTCTGGC
 AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG
 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
 3422 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTTCCTCCATGCCCCCCTGGAGGGGGAGCCT
 ACGGGGGGGCTGAGGCTGCGACTCAGGATAAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
 ^
 3437 EAM11051,
 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
 3482 GGGGATCCGGATCTTAGCGACGGGTGATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG
 CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGTTTGCGCCTC
 ^ ^ ^
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,
 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
 3542 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

3589 DRA3, 3600 SAC2,

AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
3602 GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1.

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGCTCTTCTTTCAGTGTAAACTG

3681 DRA3,

ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCTCCAATTTTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTCCGG

3875 AAT2, 3890 BGLI,

ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
3902 GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
4022 CCAGCTCGTCTCATCGTGTCCCCGATCTGGGCGTGCGGTGTGCGAAAAGATGGCTTTG
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
TACGACGTGGTTACAAAGCTCCCCTTGCCGTGATGGGAAGCTCCTACGGATTCCAATAC
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAGTCCAAGAAACCCCAATG
AGTGGTCCTGTCGCCCACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
GGGTTCCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

4893 BGL2,

4954 NCOI,

5015 SPHI, 5035 KPNI,

5064 APAI, 5091 BALI,

5113 NDEI,

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5240 DRA3,

5295 PSTI,

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
5402 CCGGGTGGCGGT CAGATCGTTGGTGGAGTTTACTTGTTGCCGCGCAGGGGCCCTAGATTG
GGCCCCACGCCAGTCTAGCAACCACTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC

5449 APAI,

GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
5462 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCGAACCTCGAGGTAGACGTACGCCT
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
5522 ATCCCCAAGGCTCGTCGGCCCCAGGGCAGGACCTGGGCTCAGCCCGGTACCTTGGCCC
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
5582 CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCGTGGCTCTCGG
GAGATACCGTTACTCCCGACGCCCCACCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp
5642 CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCTCTT
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA

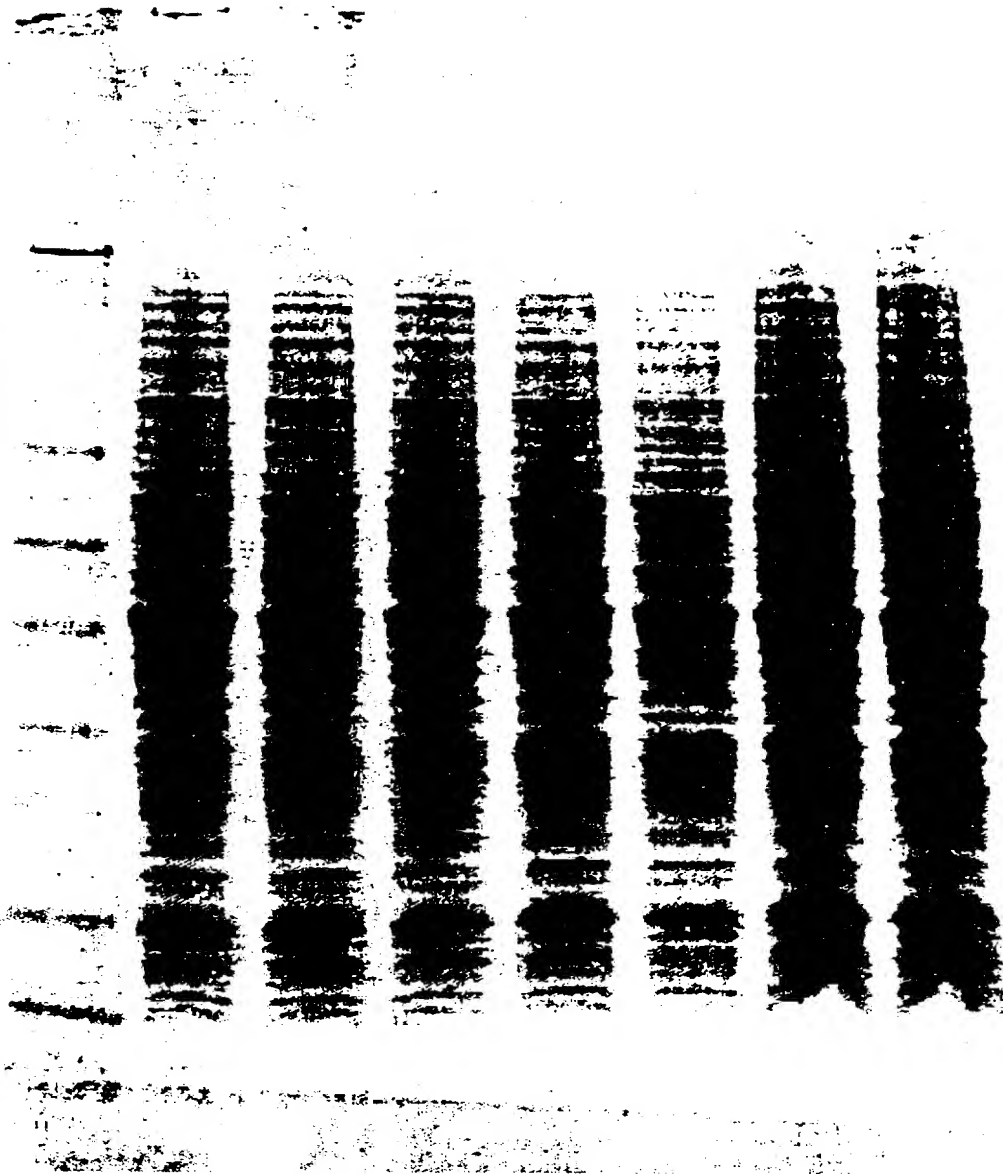
5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,

5762 GlyGlyAlaAlaArgAlaOC AM
GGAGGCGCTGCCAGGGCCTAATAGTCGAC
CCTCCGCGACGGTCCCGGATTATCAGCTG

5785 SALI,

Variable	Mean	SD	Min	Max
Age	35.2	12.5	18	65
Gender	Male	0.75	0	1
Marital Status	Married	0.65	0	1
Education	High School	0.35	0	1
Occupation	Unemployed	0.45	0	1
Income	\$15,000	\$10,000	\$0	\$50,000
Health Status	Good	0.85	0	1
Smoking	Non-smoker	0.70	0	1
Alcohol	Non-drinker	0.60	0	1
Exercise	Regular	0.55	0	1
Stress	Low	0.40	0	1
Sleep	7-8 hours	1.5	4	10
Diet	Healthy	0.65	0	1
Family Size	2-3	1.0	1	5
Neighborhood	Urban	0.50	0	1
Transportation	Car	0.80	0	1
Insurance	Health	0.90	0	1
Employment	Full-time	0.70	0	1
Unemployment	Unemployed	0.30	0	1
Wage	\$15,000	\$10,000	\$0	\$50,000
Hours	40	10	20	60
Benefits	Yes	0.85	0	1
Job Satisfaction	High	0.60	0	1
Work-Life Balance	Good	0.75	0	1
Career Development	Yes	0.50	0	1
Supervisor Rating	4.0	0.5	1.0	5.0
Peer Rating	4.0	0.5	1.0	5.0
Self-Rating	4.0	0.5	1.0	5.0
Overall Rating	4.0	0.5	1.0	5.0

FIGURE 23



002277 52479 4.2000